

STIC-Biotech/ChemLib

164529

From: Whiteman, Brian  
Sent: Wednesday, August 31, 2005 4:20 PM  
To: STIC-Biotech/ChemLib  
Subject: seq search

09/475,704 Barnett et al. 12/30/99

SEQ ID NO; 3 and 4  
Please search protein and NA databases  
1) the issued and published US application databases

Thank you,

Brian Whiteman  
Remsen, 2D14  
mail box 2C18  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
(571) 272-0764

8/31/05

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: 9/1/05  
Date Completed: 9/2/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: 2+ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: DS  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**



QY 241 ACCGTGGCCACCCTGACTGCGTGACGAGAAAGTGAAGTCCGCGACACCAAGAGGCC 300  
DB 247 ACCGTGGCCACCCTGACTGCGTGACGAGAAAGTGAAGTCCGCGACACCAAGAGGCC 306  
QY 301 CTGGACAAGATCGAGAGAGACGAAACAAGTGCACAGAAAGATCCAGAGGCCAGGCC 360  
DB 307 CTGGACAAGATCGAGAGAGACGAAACAAGTCCAGAAAGAGGCCAGAGGCCAGGCC 366  
QY 361 GCGG-----ACAAGGCAAGTGAAGCAGAACTAACCCCTGTCGAGAAACCTG 408  
DB 367 GCGGCGGCGGACCGGCAACGACGAGTGAAGCAAACTAACCCCTGTCGAGAAACCTG 426  
QY 409 CAGGGCCAGATGATGACCAAGCCATCAGCCCGGACCCCTGTAACGCTGAGTGAAGTGA 468  
DB 427 CAGGGCCAGATGATGACCAAGCCATCAGCCCGGACCCCTGTAACGCTGAGTGAAGTGA 486  
QY 469 ATCGAAGAAAGGCTTCAGCCCGGAGTGAATCCCATGTTCAACCGCCCTGAGCGAGGCC 528  
DB 487 GTGGAAGAAAGGCTTCAGCCCGGAGTGAATCCCATGTTCAACCGCCCTGAGCGAGGCC 546  
QY 529 GCGACCCCGGACCGTGAACACGATGTTGAACACCGTGGGGGCGGACGAGAGCGCGCATG 588  
DB 547 GCGACCCCGGACCGTGAACACGATGTTGAACACCGTGGGGGCGGACGAGAGCGCGCATG 606  
QY 589 CAGATGCTGAAGAGACCATCAACGAGAGCGCGGAGTGGGACCGCTGACACCCCGTG 648  
DB 607 CAGATGCTGAAGAGACCATCAACGAGAGCGCGGAGTGGGACCGCGTGAACCCCGTG 666  
QY 649 CAGCGCGGCGCCATCGCCCGGCGGAGTGGGAGCCCGGCGGACGACCATCGCGGCG 708  
DB 667 CAGCGCGGCGCCATCGCCCGGCGGAGTGGGAGCCCGGCGGACGACCATCGCGGCG 726  
QY 709 ACCACGAGACCCCTGAGAGAGCATGCTGGATGACGAGCAACCCCGCATCCCGCGT 768  
DB 727 ACCACGAGACCCCTGAGAGAGCATGCTGGATGACGAGCAACCCCGCATCCCGCGT 786  
QY 769 GCGCATCTACAGAGCGGTGATCATCTGGGCTGGAACAGATCTGGGAGTGAACG 828  
DB 787 GCGCATCTACAGAGCGGTGATCATCTGGGCTGGAACAGATCTGGGAGTGAACG 846  
QY 829 CCGGTGAGCATCTGGAACATCAAGCAGGCGCCCAAGAGCCCTTCGCGCATGCTGGAGC 888  
DB 847 CCGGTGAGCATCTGGAACATCAAGCAGGCGCCCAAGAGCCCTTCGCGCATGCTGGAGC 906  
QY 889 GCGTCTTCAAGACCCCTGCGCGCGGAGAGACCCAGAGGTGAAGAACTGAGTGAAC 948  
DB 907 GCGTCTTCAAGACCCCTGCGCGCGGAGAGACCCAGAGGTGAAGAACTGAGTGAAC 966  
QY 949 GACACCTGCTGATGAGAACGCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGC 1008  
DB 967 GACACCTGCTGATGAGAACGCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGC 1026  
QY 1009 CCGCGCGCGACCTGAGAGAGATGATGATCCGCTGCGAGGCGGTGGCGCGCCGACGAC 1068  
DB 1027 CCGCGCGCGACCTGAGAGAGATGATGATCCGCTGCGAGGCGGTGGCGCGCCGACGAC 1086  
QY 1069 AAGGCGCGGATGCTGCGCGAGGCGATGAGCAGGCGCAACA-----CAGCGTGAATGATG 1122  
DB 1087 AAGGCGCGGATGCTGCGCGAGGCGATGAGCAGGCGTGAACCCGCGCATCATGATG 1146  
QY 1123 CAGAAAGCACTTCAAGAGGCCCGCGCGCATGCTGAAGTCTTCAACTGCGCGCAAGAG 1182  
DB 1147 CAGCGCGGCAACTTCCGCAACGAGCGGAAGACCGTCAAGTGTCTTCAACTGCGCGCAAGAG 1206  
QY 1183 GCGCATGCTGCGCGCACTGCGCGCGCGCGCGCGCAAGAGGCGTCTGGAAGTGGCGAAG 1242  
DB 1207 GCGCATGCTGCGCGCACTGCGCGCGCGCGCGCGCAAGAGGCGTCTGCGCGCGCGCG 1266  
QY 1243 GAGGGCCACAGATGAAGAGCTGCAACGAGCGCGAGCGCAACTTCTGCGCGCAAGATCTG 1302  
DB 1267 GAGGGCCACAGATGAAGAGCTGCAACGAGCGCGAGCGCAACTTCTGCGCGCAAGATCTG 1326  
QY 1303 CCGAGCCACAGAGCGCGCGCGCGCAACTTCTGCAAGAGCGCGCGCGCGCGCGCGCG 1362

DB 1327 CCCAGCTCAAGAGGCGCGCGCGCAACTTCTGAGAGACCGCGCGGACCGACCGCGCG 1386  
QY 1363 CCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCAAGAGAGAG--- 1413  
DB 1387 CCGAGAGAGAGCTTCCGCTTGGCGAGAGAACCAACCCCGACGCAAGAGAGAGAGCGCC 1446  
QY 1414 ---AGCAAGAGACCGGAGACCCCTGACAGCTGAAAGAGCTGTTGCGCAAGACCCCTG 1470  
DB 1447 ATCGACAAAGAGCTTACCCCTGACAGCCTTGGCGAGCTGTTGCGCAAGACCCGACG 1506  
QY 1471 AGCCAGTAA 1479  
DB 1507 AGCCAGTAA 1515  
  
RESULT 2  
US-09-475-515-75  
/ Sequence 75, Application US/09475515A  
/ Patent No. 6602705  
/ GENERAL INFORMATION:  
/ APPLICANT: BARNETT, Susan  
/ APPLICANT: ZOR MEDEDE, Jan  
/ APPLICANT: SRIVASTAVA, Indresh  
/ APPLICANT: LIAN, Ying  
/ APPLICANT: HARTOG, Karin  
/ APPLICANT: LIU, Hong  
/ APPLICANT: GREER, Catherine  
/ APPLICANT: SEIBY, Mark  
/ APPLICANT: WALKER, Christopher  
/ TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
/ FILE REFERENCE: 1621.002  
/ CURRENT APPLICATION NUMBER: US/09/475,515A  
/ CURRENT FILING DATE: 1999-12-30  
/ NUMBER OF SEQ ID NOS: 90  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 75  
/ LENGTH: 4472  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence:  
/ OTHER INFORMATION: gp160.modUS4.delV1/V2.gag.modSF2  
US-09-475-515-75  
  
Query Match 81.0%; Score 1197.8; DB 4; Length 4472;  
Best Local Similarity 89.3%; Pred. No. 9.3e-177;  
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;  
  
QY 1 ATGGGCGCGCGCGGACGATCTGCGCGCGGCAAGCTGAGCGCCTGGGAGCGCATCCGC 60  
DB 2939 ATGGGCGCGCGCGGACGATCTGCGCGCGGCAAGCTGAGCGCCTGGGAGCGCATCCG 2998  
QY 61 CTGCGCGCGCGCGGCAAGAGTCTACATGATGAAGCACTGTGTGTGGCGGACGCGCGAG 120  
DB 2939 CTGCGCGCGCGCGGCAAGAGTCTACATGATGAAGCACTGTGTGTGGCGGACGCGCGAG 3058  
QY 121 CTGGAAGAGTTCGCTTAACCCCGGCTGCTGAGAGACGAGGAGGCTGCAAGCATC 180  
DB 3059 CTGGAAGAGTTCGCTTAACCCCGGCTGCTGAGAGACGAGGAGGCTGCGCGCGCATTC 3118  
QY 181 ATCGCGAGCTGACACCGCGCCTGCGAGCGGCGAGCGGAGCTGAAAGAGCTGTTCAAC 240  
DB 3119 CTGGGCGAGCTGAGCGCGCCTGCGAGCGGCGAGCGGAGCTGCGCGCTGTTCAAC 3178  
QY 241 ACCGTGGCCACCCTGACTGCGTGACGAGAAAGATCGAGTCCGCGACCAAGAGAGGCC 300  
DB 3179 ACCGTGGCCACCCTGACTGCGTGACGAGAAAGATCGAGTCCGCGACCAAGAGAGGCC 3238  
QY 301 CTGGAAGAGATCGAGAGAGAGCAAGCAAGTGCAGCAAGAAAGATTCAGCAGGCGCGAGGCC 360  
DB 3239 CTGGAAGAGATCGAGAGAGAGCAAGCAAGTGCAGCAAGAAAGAGGCCAGCAGGCGCGCG 3298



```

QY 361 GCGG-----ACAGGGCAAGTGAAGCCAGAACTAACCCCAATCGTGAGAACTCTG 408
    |||||
Db 3299 GCGGCGGGCAAGCCGCAACAGCAAGCCAGGTGAGCCAGAACTAACCCCAATCGTGAGAACTCTG 3358
    |||||
QY 409 CAGGGCCAGATGTGTGACCAAGGCAATCAGCCCCCGACCTGTGAACGCTGTGTAAGTGTG 468
    |||||
Db 3359 CAGGGCCAGATGTGTGACCAAGGCAATCAGCCCCCGACCTGTGAACGCTGTGTAAGTGTG 3418
    |||||
QY 469 ATCGAGAGAAAGGCTTTCAGCCCCCGAGGTGATCCCATGTTCAACGCGCTTGAGCGAGGGC 528
    |||||
Db 3419 GTGAGAGAAAGGCTTTCAGCCCCCGAGGTGATCCCATGTTCAACGCGCTTGAGCGAGGGC 3478
    |||||
QY 529 GCCACCCCGCAGGACCTGAAACAAGTGTGAACACCGTGGGGGGCCACAGAGCGCGCAGT 588
    |||||
Db 3479 GCCACCCCGCAGGACCTGAAACAAGTGTGAACACCGTGGGGGGCCACAGAGCGCGCAGT 3538
    |||||
QY 589 CAGATGCTGAAGAGCAACATCAACGAGAGGCGCGGAGTGGGACCGCGTGCAACCCCGTG 648
    |||||
Db 3539 CAGATGCTGAAGAGCAACATCAACGAGAGGCGCGGAGTGGGACCGCGTGCAACCCCGTG 3598
    |||||
QY 649 CACGCGGCGCCCATCGCCCCCGGCGAGATGCGGAGCCCGCGGCGAGCAATCGCGGCG 708
    |||||
Db 3599 CACGCGGCGCCCATCGCCCCCGGCGAGATGCGGAGCCCGCGGCGAGCAATCGCGGCG 3658
    |||||
QY 709 ACCACGAGACCCCTGAGAGAGAGATGCGCTGTGATGACAGCAACCCCGCCATCCCGCGT 768
    |||||
Db 3659 ACCACGAGACCCCTGAGAGAGAGATGCGCTGTGATGATCAACAACCCCGCCATCCCGCGT 3718
    |||||
QY 769 GCGCATCTACAGAGCGGTGATCATCTGTGCGCTGTAACAAGATCTGTGCGATGTACAGC 828
    |||||
Db 3719 GCGCATCTACAGAGCGGTGATCATCTGTGCGCTGTAACAAGATCTGTGCGATGTACAGC 3778
    |||||
QY 829 CCGGTGAGCATCTGTGACATCAAGCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGAGAC 888
    |||||
Db 3779 CCGGTGAGCATCTGTGACATCAAGCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGAGAC 3838
    |||||
QY 889 CGCTTTCAGAGCCCTGTGGCGCGCGAGAGCAACCCAGAGGTGAAGAACTGAGTGAAC 948
    |||||
Db 3839 CGCTTTCAGAGCCCTGTGGCGCGCGAGAGCAACCCAGAGGTGAAGAACTGAGTGAAC 3898
    |||||
QY 949 GACACCTCTGTGTGACAGACGCCCAACCCCGACTGTGAAGAACATCTGTGCGCGCTCGGC 1008
    |||||
Db 3899 GACACCTCTGTGTGACAGACGCCCAACCCCGACTGTGAAGAACATCTGTGCGCGCTCGGC 3958
    |||||
QY 1009 CCGCGCGCGCAGCTGTGAGAGATGATGACCGCTTGCAGAGCGCGTGGGCGGCCCCAGCCAC 1068
    |||||
Db 3959 CCGCGCGCGCAGCTGTGAGAGATGATGACCGCTTGCAGAGCGCGTGGGCGGCCCCAGCCAC 4018
    |||||
QY 1069 AAGGCGCGCGGTGTGCGCCAGAGCGATGAGACCAAGCCCAACA-----CAAGCTGATGATG 1122
    |||||
Db 4019 AAGGCGCGCGGTGTGCGCCAGAGCGATGAGACCAAGCCCAACA-----CAAGCTGATGATG 4078
    |||||
QY 1123 CAGAAAGAGCACTTCAAGGGGCCCCCGCGCATGTCAAGTGTCTTCAACTGCGGCAAGAGAG 1182
    |||||
Db 4079 CAGCGCGGCAACTTCCGCAACCAAGGAAAGACCGTGAATGTCTTCAACTGCGGCAAGAGAG 4138
    |||||
QY 1183 GGCACATGCGCGCAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGAAGTGTGCGGCAAG 1242
    |||||
Db 4139 GGCACACCGCGAGAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGCGGCGCGCGCGC 4198
    |||||
QY 1243 GAGGGGCAACAGATGAAGAGCTGCAACCGAGCGCGCAAGCCCACTTCTTGGGCAAGATGTG 1302
    |||||
Db 4199 GAGGGGCAACAGATGAAGAGCTGCAACCGAGCGCGCAAGCCCACTTCTTGGGCAAGATGTG 4258
    |||||
QY 1303 CCGAGCGCAAGAGGCGCGCGCGCAACTTCTTGAAGAGCGCGCGCGAGCGCAACCGCGCGC 1362
    |||||
Db 4259 CCGAGCTACAGAGGCGCGCGCGCAACTTCTTGAAGAGCGCGCGCGAGCGCAACCGCGCGC 4318
    |||||
QY 1363 CCGCGCGAGAGCTTCCGCTTCT-----GAGAGACCAACCCCGCGCGCAAGAGAGAG--- 1413
    |||||
Db 4319 CCGGAGAGAGCTTCCGCTTCTCGCGAGAGAGACCAACCCCGCGCGCAAGAGAGAGAGGCC 4378
    |||||

```

```

QY 1414 ---AGCAAGAGACCGGAGAGCCCTGACCAAGCTGAAGAGCTGTGCGAGAGAGCCCGCTG 1470
    |||||
Db 4379 ATCGACAGAGAGCTGTACCCCTTGACAGACTTCCGAGAGCTGTGTGCGAGAGAGCCCGCAGC 4438
    |||||
QY 1471 AGCCAGTAA 1479
    |||||
Db 4439 AGCCAGTAA 4447
    |||||

RESULT 3
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 76
; LENGTH: 4608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modsfp162.delv2.gag.modsfp2
US-09-475-515-76

Query Match      81.0%; Score 1197.8; DB 4; Length 4608;
Best Local Similarity 89.3%; Pred. No. 9,3e-177;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCGCAGATCTGTGCGCGCGGCAAGCTGAGCGCTGTGAGGCGATCCGC 60
    |||||
Db 3075 ATGGGCGCGCGCGCGCAGCGTGTGAGCGCGCGCGCTGACATGTGTGTGGGCCAGCGCGAG 3134
    |||||
QY 61 CTGGGCGCGCGCGGCAAGTGTCTAATGATGAAGCACTGTGTGTGGGCCAGCGCGAG 120
    |||||
Db 3135 CTGGGCGCGCGCGGCAAGTGTCTAATGATGAAGCACTGTGTGTGGGCCAGCGCGAG 3194
    |||||
QY 121 CTGGAAGATTGCGCCTTAACCCCGCTGTGTGAGACCAAGAGAGGCTGTCAAGCAGATC 180
    |||||
Db 3195 CTGGAAGCCTTGCGCTGTAAACCCCGCTGTGTGAGACCAAGAGAGGCTGTGCGCAGATC 3254
    |||||
QY 181 ATCCGCAAGCTGCAACCCCGCTGTGCAACCGGCAAGAGAGCTGAAGAGCTGTTCAAC 240
    |||||
Db 3255 CTGGGCGAGCTGCAACCCCGCTGTGCAACCGGCAAGAGAGCTGCGAGCTGTATCAAC 3314
    |||||
QY 241 ACCGCGGCAACCTGTACTGTGTGTGACAGAAAGATCGAGGTCCGCGACCAAGAGAGGCC 300
    |||||
Db 3315 ACCGCGGCAACCTGTACTGTGTGTGACAGAGGATCGATGAGTCAAGAGACCAAGAGAGGCC 3374
    |||||
QY 301 CTGCACAAGATCGAGAGAGAGAGCAAGTGTCCAGCAAGAGATTCACAGGCGCGAGGCC 360
    |||||
Db 3375 CTGGAAGAGATCGAGAGAGAGAGCAAGTGTCAAGAGAGAGGCCACAGGCGCGCGC 3434
    |||||
QY 361 GCGG-----ACAGGGCAAGGTGAGGAGCCCAACTACCTCATGTGTGCAAGAACTTG 408
    |||||
Db 3435 GCGGCGGCAACCGGCAAGAGCCAGGTGAGCCAGAACTACCTCATGTGTGCAAGAACTTG 3494
    |||||
QY 409 CAGGCGAGATGTGACCAAGGCAATCAGCCCGCGCACTTGAAGCGCTGTGGGTGAAGGTG 468
    |||||

```

Db	3455	CAGGGCCAGATGTGTGACACCAAGGCCATCAGCCCCCGGACCCCTGAAACGCTGGGTGAAGGTG	355
Qy	469	ATCGAGAGAAAGGCTTTCAGCCCCCGAGGTGATCCCATGTTCAACGGCTTGAGCAGGAGGC	528
Db	3555	GTGGAGGAGAAAGGCTTTCAGCCCCCGAGGTGATCCCATGTTTCAGCGCCTTGAGCAAGGGC	3614
Qy	529	GGCACCCTCCCGAGACCTGAAACAACGATGTTTGAACACCGTGGGGGGGCAACAGGCCGCTATG	588
Db	3615	GGCAACCCCCCGAGACCTGAAACAACGATGTTTGAACACCGTGGGGGGGCAACAGGCCGCTATG	3674
Qy	589	CAGATGTGTAAGAGACCATCAACAAGAGAGGCGCGGATGGGGAACGGGTGCAACCCGCTG	648
Db	3675	CAGATGTGTAAGAGACCATCAACAAGAGAGGCGCGGATGGGGAACCGGTGCAACCCGCTG	3733
Qy	649	CACGCGGAGCCCATGCGCCCGCGGCGAGATGCGCGAGCCCGCGGAGGACATGCGCGGC	708
Db	3735	CACGCGGAGCCCATGCGCCCGCGGCGAGATGCGCGAGCCCGCGGAGGACATGCGCGGC	379
Qy	709	ACCACCAAGACCTCTGCAGAGAGCAGATGCGCTGTGATGACCAAGCAACCCCTCATCCCGTG	768
Db	3795	ACCACCAAGACCTCTGCAGAGAGCAGATGCGCTGTGATGACCAACACCCCTCATCCCGCTG	385
Qy	769	GGCGCATCTTCAAGCGGTGATCATCTCGGGGCTGGAACAAGATGTGCGGATGTACAGC	828
Db	3855	GGCGAGATCTTCAAGCGGTGATCATCTCGGGCTGGAACAAGATGTGCGGATGTACAGC	3914
Qy	829	CCCGTGACATCTTGGACATCAAGCAGGAGGCCCAAGAGACCTTTCGCGACTTACGTGAGC	888
Db	3915	CCCAACCAACATCTGTGACATCGCGCAGAGGGCCCAAGAGACCTTTCGCGACTTACGTGAGC	3974
Qy	889	CGCTTCTTCAAGACCTCTGCGCGCGCGAGCAAGACCCCAAGAGGTGAAGATCTGATGACC	948
Db	3975	CGCTTCTTCAAGACCTCTGCGCGCTGTAGAGAGGCAAGCAAGTGAAGAACTGTGATGACC	4034
Qy	949	GACACCCGTGTGTGTGACAAAGGCCAACCCCGATCTGACAGACCATCTGCGCTTGGGC	1008
Db	4035	GAGACCTCTGTGTGTGACAAAGGCCAACCCCGATCTGACAGACCATCTTGAAGCTTGGGC	4094
Qy	1009	CCCGGCGCGAGCTTGGAGAGATGTATGACCGCTCTGCAGAGGGGTGGGGCGGCCCAAGCAC	1068
Db	4095	CCCGGCGCGACCTCTGGAGAGATGTATGACCGCTCTGCAGAGGGGTGGGGCGGCCCAAGCAC	4155
Qy	1069	AAAGCCGCGGTGTGTGGCCGAGGCGCATGAGCCAGGCGAAACA-----CGAGCTGATGTG	11224
Db	4155	AAAGCCGCGGTGTGTGGCCGAGGCGCATGAGCCAGGCGAAACC CGGACATCATGTATG	4214
Qy	1123	CAGAAAGACATCTCAAGGCCCCCGGCGCATCTGTCAAGTGTTTCAATCTGGCGAAAGAG	11824
Db	4215	CAGCGCGGCAACTTCCGACCACTCAGGGGAAAGCCGTCAAGTGTTTCAATCTGGCGAAAGAG	4274
Qy	1183	GGCCACATGTGCGCGGAACTGTGCGCGCGCCCCCGCAAGAAAGGGCTGTGGAGTGGCGCAAG	12424
Db	4275	GGCCACATGTGCGCGGAACTGTGCGCGCGCCCCCGCAAGAAAGGGCTGTGGAGTGGCGCGC	4333
Qy	1243	GAGGGCCACCAAGATGAAGACTGCAACGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGG	13024
Db	4335	GAGGGCCACCAAGATGAAGACTGCAACGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGG	4394
Qy	1303	CCCAAGCCCAAGAGGCGCGCGCGGACAATTCTGTGCAAGCGCGCCGAGGCCACCGCCCCC	13674
Db	4395	CCCAAGCTTCAAGAGGCGCGCGCGGACAATTCTGTGCAAGCGCGCGCCGAGGCCACCGCCCCC	4454
Qy	1363	CCCGCGAGAGTTCGCGCTTC-----GAGGAGACACCCCGCGCAGAAAGCAGAG--	1413
Db	4455	CCCGAGAGAGTTCGCGCTTCGCGCGAGGAGAAACAACCCGACGCAAGAGCAGAGCGCC	4514
Qy	1414	---AGCAAGACCGCGAGACCTTGAACCAAGCTTGAAGAGCTTGTGCGCAACGACCCCTG	14707
Db	4515	ATGCAAGAGAGCTTACCCCTGTACAGCCGCTGGCAGAGCTGTGTTGGCAAGAACCCAGC	4574
Qy	1471	AGCCAGTAA 1479	
Db	4575	AGCCAGTAA 4583	

```

/ RESULT 4
/ US-09-475-515-74
/ Sequence 74, Application US/09475515A
/ Patent No. 6602705
/ GENERAL INFORMATION:
/ APPLICANT: BARNETT, Susan
/ APPLICANT: ZUR MEGEDE, Jan
/ APPLICANT: SRIVASTAVA, Indresh
/ APPLICANT: LIAN, Ying
/ APPLICANT: HARTOG, Karin
/ APPLICANT: LIU, Hong
/ APPLICANT: GREER, Catherine
/ APPLICANT: SELBY, Mark
/ APPLICANT: WALKER, Christopher
/ TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
/ TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
/ FILE REFERENCE: 1621.002
/ CURRENT APPLICATION NUMBER: US/09/475,515A
/ CURRENT FILING DATE: 1999-12-30
/ NUMBER OF SEQ ID NOS: 90
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 74
/ LENGTH: 4689
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:
/ OTHER INFORMATION: gp160.modsf162.g9g.modsf2
/ US-09-475-515-74

```

Query Match	81.0%	Score 1197.8	DB 4	Length 4689
Best Local Similarity	89.3%	Pred. No. 9.3e-177		
Matches 1347	Conservative	0	Mismatches 132	Indels 30
			Gaps	4
QY	1	ATGGGCGCGCCGAGCAGCATCTGTGGGGCGGCGGCAAGCTGAGACGCTGTGGAGCGCATCCG	60	
DB	3156	ATGGGCGCGCCGCGCAGCGTCTTACGCGCGGCGGCGAGCTGGACAACTGGGAGAAAGTCCCG	3215	
QY	61	CTGGCGCCCGCGGCGAAGAGTGTACATGATGAAGCATCTGTGTGGGCGAGCCGCGAG	120	
DB	3216	CTGGCGCCCGCGGCGAAGAAAGTAACTGAAGCATCTGTGTGGGCGAGCCGCGAG	3275	
QY	121	CTGGAGAAAGTTCGCCCTGGAACCCCGCGCTGTGTGAGAACACCGAGGCTTCAAGCATC	180	
DB	3276	CTGGAGCGCTTCGCCGTAACCCCGCGCTGTGTGAGAACACCGAGGCTTCCGCTCAAGTC	3335	
QY	181	ATCGCGCAGCTGCACCCCGCGCTGTGACAGCCGACGAGAGCTGAAGAGCTGTTCAC	240	
DB	3336	CTGGGCGCAGCTGCACCCCGCGCTGTGACAGCCGACGAGAGCTGTCCGAGCTGTACAC	3395	
QY	241	ACCGTGCGCACCTGTACTGTGCTGACAGAGAATCGAGGTCCGCGACCAAGAGAGCC	300	
DB	3396	ACCGTGCGCACCTGTACTGTGCTGACCAAGAGCATCGTCAAGAGACACCAAGAGAGCC	3455	
QY	301	CTGGACAAAGATTCGAGAGGAGACAAACAAAGTCCAGAGAAATTCAGACAGCCGAGGCC	360	
DB	3456	CTGGAGAAAGATTCGAGAGGAGACAAACAAAGTCCAGAGAAAGGCCCGACAGCGCGCCGC	3515	
QY	361	GCCG-----ACAAAGGAGAGGTGAGCAGAACTACCCCATGTGTACAGAACTGTG	408	
DB	3516	GCCGCGCGACCGGCAACGACGACAGGTGAGCCAGAACTACCCCATGTGTACAGAACTGTG	3575	
QY	409	CAGGCGCAGATGTGACACAGGCGCATAGCCCCGCAACCTTGAAACGCTTGAGTGAAGTG	468	
DB	3576	CAGGCGCAGATGTGACACAGGCGCATAGCCCCGCAACCTTGAAACGCTTGAGTGAAGTG	3635	
QY	469	ATCGAGAGAGAGGCTTCAGGCGCGGAGGTATCCCATGTTCAACGCCCTAGAGGAGAGGC	528	
DB	3636	GTGAGAGAGAGGCTTCAGGCGCGGAGGTATCCCATGTTCAAGGCCCTTAGAGCAAGGCGC	3695	
QY	529	GCAACCCCCCAGACCTGAAACCATGATTGTAACACGTTGAGGCGGCGACAGGCGGCATGTG	588	

Db 3696 GCGACCCCCGACGAGACTGAAACAAGATGTTGAAACCCGTGGGGCGGCAACAGCGCGCAATG 3755

QY 589 CAGATGCTGAAGAGACAATCAACGAGAGGCGCGCGAGTGGAGCCGGTGGCAACCCCGTG 648

Db 3756 CAGATGCTGAAGAGAGACCATCAACGAGAGGGCGCGCGAGTGGAGCCGGTGGCAACCCCGTG 3815

QY 649 CACCGCGGGCCCCCATGCGCCCCCGGCGCGAGATGGCGAGGCCCGCGGACGAGCATGGCGGG 708

Db 3816 CACGCGGGCCCCCATGCGCCCCCGGCGCGAGATGGCGAGGCCCGCGGACGAGCATGGCGGGC 3875

QY 709 ACCACCAACACCCCTGCAGAGAGAGATGGCTGGATGACACAGCAACCCCCCATCCCGGTG 768

Db 3876 ACCACCAACACCCCTGCAGAGAGAGATGGCTGGATGACACACACACCCCCCATCCCGGTG 3935

QY 769 GCGGACATCTACAGCGGGTGGATCATCTGGGCGTTGAAACAAGATGTCGATGTACAGC 828

Db 3936 GCGGAGATCTACAGCGGGTGGATCATCTGGGCGTTGAAACAAGATGTCGATGTACAGC 3995

QY 829 CCGGTGAGCATCTTGGACATCAAGAGAGGCCCCCAAGAGCCCTTCCGGACTAGGTGAC 888

Db 3996 CCGACACACATCTTGGACATCGCACGAGGGCCCCCAAGAGGCCCTTCCGGACTAGGTGAC 4055

QY 889 CGCTCTTCAAGAACCCTGCGCGCGCGAGAGAGACCCAGAGGGTGAACAATCGATGAGAC 948

Db 4056 CGCTTCTACAGAACCTTCCGGCTGGAGAGGCGACCGACGAGACGTGAACATCGATGAGACC 4115

QY 949 GACACCTGTGTTGTCAGAGAGCCCAACCCCGACTGCAAGACATCTTCCGCTCGGC 1008

Db 4116 GAGACCTGTGTTGTCAGAGAGCCCAACCCCGACTGCAAGACATCTTGAAGGCTCTCGGC 4175

QY 1009 CCGCGGCGCGACCTGGAGAGATGATGACCCGCTTCGAGGGCGTGGGCGGCCGAC 1068

Db 4176 CCGCGGCGCGACCTGGAGAGATGATGACCCGCTTCGAGGGCGTGGGCGGCCGAC 4235

QY 1069 AAGGCGCGCGTGGCGAGGGCCATGAGCCAGGCAACA-----CAAGGTGATGATG 1122

Db 4236 AAGGCGCGCGTGGCGAGGGCCATGAGCCAGGCGAGAACCCGCGACCATATGATG 4295

QY 1123 CAGAAAGACCACTTCAAGAGGCCCCCGGCGCATCTGCAAGTCTTCAACTGCGGCAAGAG 1182

Db 4296 CAGGCGGCGCACTTCCGACACCAAGCGAAGACCGTCAAGTCTTCAACTGCGGCAAGAG 4355

QY 1183 GCGCACATCGCGCGCACTGCGCGGCCCGCCCGCAAGAGGGCTCTGGAAGTGGCGCAAG 1242

Db 4356 GCGCACACCGCGCAAGAACTGCGCGGCCCGCCCGCAAGAGGGCTCTGCGCGTGGCGCGC 4415

QY 1243 GAGGGGCCCAAGATGAAGAGACTGCAACGAGGCGCCAGGCAACTTCTTGGGCAAGATCTGG 1302

Db 4416 GAGGGGCCCAAGATGAAGAGACTGCAACGAGGCGCCAGGCAACTTCTTGGGCAAGATCTGG 4475

QY 1303 CCGACGCAAGAGGGCGCGCGCGGCACTTCTTGGCAAGGCGGCCCGAGGCCACGCGCCCC 1362

Db 4476 CCGACGCAAGAGGGCGCGCGCGGCACTTCTTGGCAAGGCGGCCCGAGGCCACGCGCCCC 4535

QY 1363 CCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCGCAAGACAGAG-- 1413

Db 4536 CCGGAGGAGAGCTTCCGCTTCGCGGAGGAGAAACCAACCCCGACGAGAACAGAGAGGCC 4595

QY 1414 ---AGCAAGAGACCGGAGAGACCTTGAACAGCTTGAAGGCTGTTGGCAACGACCCCTGTG 1470

Db 4596 ATCGACAGAGAGCTGTATACCCCTGACCAAGCTTGGCGACGCTGTTCGCGCAAGACCCAGC 4655

QY 1471 AGCCAGTAA 1479

Db 4656 AGCCAGTAA 4664

RESULT 5  
US-09-475-515-73  
; Sequence 73 Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:

```

1  APPLICANT: BARNETT, Susan
2  APPLICANT: ZUR MEGEDE, Jan
3  APPLICANT: SRIVASTAVA, Indresh
4  APPLICANT: LIAN, Ying
5  APPLICANT: HARTOG, Karin
6  APPLICANT: LIU, Hong
7  APPLICANT: GREER, Catherine
8  APPLICANT: SELBY, Mark
9  APPLICANT: WALKER, Christopher
10 TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
11 TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
12 FILE REFERENCE: 1621.002
13 CURRENT APPLICATION NUMBER: US/09/475.515A
14 CURRENT FILING DATE: 1999-12-30
15 NUMBER OF SEQ ID NOS: 90
16 SOFTWARE: PatentIn Ver. 2.0
17 SEQ ID NO 73
18
19 LENGTH: 4766
20 TYPE: DNA
21 ORGANISM: Artificial Sequence
22 FEATURE:
23 OTHER INFORMATION: Description of Artificial Sequence:
24
25 US-09-475-515-73

```

Query Match	81.0%	Score 1197.8	DB 4	Length 4766	
Best Local Similarity	89.3%	Pred. No. 9.3e-177			
Matches 1347	Conservative 0	Mismatch 132	Indels 30	Gaps 4	
Qy	1	ATGGGCGCCCGCGCCAGCATCTTGCGCGCGCGCAAGCTGGA	CGCGCTGGGA	CGCATCCG	60
Db	3233	ATGGCGCGCGCGCGCCAGCGTCTGAGCGCGCGCGCTGGAC	CAAGTGGGA	GAAGATCCGC	3292
Qy	61	CTGGCGCCCGCGCGCAAGAGTGTATCATATGAAGCACTTG	GTGTGGGCC	CAAGCCCGG	120
Db	3293	CTGGCGCCCGCGCGCAAGAAAGTATCAAGCTGAAGCATCT	GTGTGGGCC	CAAGCCCGG	3352
Qy	121	CTGGAGAAGTTCGCGCTGGAACCCCGCGCTGTGAGAC	CAAGAGGCTG	CAAGCATG	180
Db	3353	CTGAGAGGCGTTTCGCGGTGAACCCCGCGCTGTGAGAC	CAAGAGGCTG	CGCCATG	3412
Qy	181	ATCCGCCAGCTGCAACCCCGCTTCGAGCCGCGAGGCTGA	AGAGCTTTCA	C	240
Db	3413	CTGGGCCAGCTGCGAGCCCGCATTCAGACCGCGAGGAG	CTGCGCATG	CAAC	3472
Qy	241	ACCGTGGCCACCCCTGTACTGCGTGCAGAGAAATCGAG	ATCCGGA	CACCAAGAGGCC	300
Db	3473	ACCGTGGCCACCCCTGTACTGCGTGCACAGCGCATCG	ACGTCAAGAC	CAACCAAGAGGCC	3532
Qy	301	CTGGACAAGATCGAGAGAGAGCAGAACAAAGTGCAG	ACAGAAAT	TCAGACAGCCGAGCC	360
Db	3533	CTGGAGAAGATCGAGAGAGAGCAGAACAAAGTCCAA	GAAGAGGCC	CCAGACAGCCGCGCC	3592
Qy	361	GGCG-----ACAAAGGGCAAGGTGAGCCAGAACTA	CCCCATGT	GTACAAAGCC	408
Db	3593	GGCGCCGCGCACCGGCAACAGCAGCAGGTGAGCCAG	AACTACCC	CATGTGTCAAGAACTG	3652
Qy	409	CAGGGCCAGATGTGTGACCAAGGCCATCAGCCCCCG	CACTTGA	CGCTTGGGTGAAGTG	468
Db	3653	CAGGGCCAGATGTGTGACCAAGGCCATCAGCCCCCG	CACTTGA	CGCTTGGGTGAAGTG	3712
Qy	469	ATCGAAGAGAAAGCGCTTCAAGCCCGGAGGTATCC	CCATGTT	CACCGCTTGAAGGAGC	528
Db	3713	GTGGAGGAGAAAGCGCTTCAAGCCCGGAGGTATCC	CCATGTT	CAACCGCTTGAAGGAGC	3772
Qy	529	GGCAACCCCGCGAGCCTGAACACGATTTTGAACA	CCGTGGCGCG	CACACAGGCCGCGCATG	588
Db	3773	GGCAACCCCGCGAGCCTGAACACGATTTTGAACA	CCGTGGCGCG	CACACAGGCCGCGCATG	3832
Qy	589	CAGATGCTGAAGACACATCAACGAGAGAGCGCC	CGATGTGG	ACCGCTGTCAACCCCGTG	648
Db	3833	CAGATGCTGAAGACACATCAACGAGAGAGCGCC	CGATGTGG	ACCGCTGTCAACCCCGTG	3892

```

QY 649 CACGCCGCGCCCATCGCCCCCGGCGAGATCGCGAGCCCGCGGCGAGCATCCGCGGC 708
| | | | |
DB 3893 CACGCCGCGCCCATCGCCCCCGGCGAGATCGCGAGCCCGCGGCGAGCATCCGCGGC 3952
| | | | |
QY 709 ACCACAGCAGCCTCGAGAGAGATCGCTTGATGACCAAGAACCCCTCCATCCCGTG 768
| | | | |
DB 3953 ACCACAGCAGCCTCGAGAGAGATCGCTTGATGACCAAGAACCCCTCCATCCCGTG 4012
| | | | |
QY 769 GCGCATATACAGAGCGGTGATCATCTCGGGCTGGAACAAGATGCGTGGATGATACGC 828
| | | | |
DB 4013 GCGCATATACAGAGCGGTGATCATCTCGGGCTGGAACAAGATGCGTGGATGATACGC 4072
| | | | |
QY 829 CCGGTGAGCATCTGAGACATCAAGAGGCGCCCAAGAGCCCTTCGCGACTACGTGAC 888
| | | | |
DB 4073 CCGGTGAGCATCTGAGACATCAAGAGGCGCCCAAGAGCCCTTCGCGACTACGTGAC 4132
| | | | |
QY 889 CGCTTCTTAAGACCTTGCGCGCCGAGAGACACCCAGAGGTGAAAGACTGATGAC 948
| | | | |
DB 4133 CGCTTCTTAAGACCTTGCGCGCCGAGAGACACCCAGAGGTGAAAGACTGATGAC 4192
| | | | |
QY 949 GACACCTGCTGTGAGAACGCAACCCGCACTGCAAGACCATCTCGCGCTCTCGGC 1008
| | | | |
DB 4193 GACACCTGCTGTGAGAACGCAACCCGCACTGCAAGACCATCTCGCGCTCTCGGC 4252
| | | | |
QY 1009 CCGGCGCGCAGCCTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGCCCGCAGCCAC 1068
| | | | |
DB 4253 CCGGCGCGCAGCCTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGCCCGCAGCCAC 4312
| | | | |
QY 1069 AAGGCGCGGTGCTGCGCGAGCGATGAGCCAGGCCAACA-----CGAGCTGATGATG 1122
| | | | |
DB 4313 AAGGCGCGGTGCTGCGCGAGCGATGAGCCAGGTGACCAACCGCGCACTCATGATG 4372
| | | | |
QY 1123 CAGAAAGCACTTCAAGGCGCCCGCGCGCATGCTCAAGTGTCTTCACTGCGGCAAGAG 1182
| | | | |
DB 4373 CAGCGCGGCACTTCCGCAACAGCGGAAGACCGTCAAGTGTCTTCACTGCGGCAAGAG 4432
| | | | |
QY 1183 GGCACATGCGCCGCAACTGCGCGCGCCCGCCCAAGAGGCTGTGAAAGTCGCGCAG 1242
| | | | |
DB 4433 GGCACATGCGCCGCAACTGCGCGCGCCCGCCCAAGAGGCTGTGAAAGTCGCGCGC 4492
| | | | |
QY 1243 GAGGCGCACATGAAAGATGTCACCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGG 1302
| | | | |
DB 4493 GAGGCGCACATGAAAGATGTCACCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGG 4552
| | | | |
QY 1303 CCGACGCAAGAGGCGCGCCGCGCACTTCTGCAAGCGCGCCGAGGCCACCGCGCCC 1362
| | | | |
DB 4553 CCGACGCAAGAGGCGCGCCGCGCACTTCTGCAAGCGCGCCGAGGCCACCGCGCCC 4612
| | | | |
QY 1363 CCGCGCGAGAGCTTCCGCTT-----GAGAGACCAACCCCGCGCAAGAGAGAG-- 1413
| | | | |
DB 4613 CCGCGAGAGAGCTTCCGCTTCCGCGAGAGAGACCAACCCCGCAAGAGAGAGAGCC 4672
| | | | |
QY 1414 ---ACCAAGAGACCGGAGACCTTGACCAAGCTGGAAGAGCTTGGGCAAGACCCCTG 1470
| | | | |
DB 4673 ATCGACAAAGAGCTGTATCCCTGACAGACCTGCGAGCTGTGGCAAGACCCAGC 4732
| | | | |
QY 1471 AGCCAGTAA 1479
| | | | |
DB 4733 AGCCAGTAA 4741
| | | | |

```

RESULT 6  
US-09-475-515-7

; Sequence 7, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Yang

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

```

; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475.515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 7
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-gag/HCV-core fusion polypeptide
US-09-475-515-7

```

Query Match 80.9%; Score 1195.8; DB 4; Length 2031;

Best Local Similarity 89.3%; Pred. No. 1.9e-176;

Matches 1345; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

```

QY 1 ATGGGCGCGCGCGCGCATCTTGCGCGCGGCAAGCTGACGCGCTGGAGCGCATCCGC 60
| | | | |
DB 7 ATGGGCGCGCGCGCGCATCTTGCGCGCGGCAAGCTGACGCGCTGGAGCGCATCCGC 66
| | | | |
QY 61 CTGGCGCGCGCGCGCAAGAGTGTACTGATGAAAGCACTGTGTGGGCCAGCGCGAG 120
| | | | |
DB 67 CTGGCGCGCGCGCGCAAGAGTGTACTGATGAAAGCACTGTGTGGGCCAGCGCGAG 126
| | | | |
QY 121 CTGGAGAGTTGCGCTGAAACCCCGCGCTGTGAGACAGAGGCGCTGCAAGCAGATC 180
| | | | |
DB 127 CTGGAGAGCTTTCGCGGTGAAACCCCGCGCTGTGAGACAGAGGCGCTGCGCAATC 186
| | | | |
QY 181 ATCGCCAGCTGACACCCCGCTGCAAGCCGCGAGGAGAGCTGAAAGCTTGTCAAC 240
| | | | |
DB 187 CTGGCGCGAGCTGACAGCCCGAGCTGCAAGCCGCGAGGAGAGCTGCGCAGCTTGAAC 246
| | | | |
QY 241 ACCGCGCGCACCGCTGTACTGCGGTGCAAGAGATCGAGGTCGCGGACCAAGAGAGCC 300
| | | | |
DB 247 ACCGCGCGCACCGCTGTACTGCGGTGCAAGAGATCGAGGTCGCGGACCAAGAGAGCC 306
| | | | |
QY 301 CTGGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCC 360
| | | | |
DB 307 CTGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCC 366
| | | | |
QY 361 GCGG-----ACAGAGGCAAGGTGAGCCAGAACTACCCCATGTGTGCAAACTTG 408
| | | | |
DB 367 GCGGCGCGCACCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCC 426
| | | | |
QY 409 CAGGCGCAGATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
| | | | |
DB 427 CAGGCGCAGATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
| | | | |
QY 469 ATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
| | | | |
DB 487 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
| | | | |
QY 529 GCGACCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
| | | | |
DB 547 GCGACCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
| | | | |
QY 589 CAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
| | | | |
DB 607 CAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
| | | | |
QY 649 CACGCCGCGCCCATCGCCCCCGGCGAGATCGCGAGCCCGCGGAGAGAGATCGCGGC 708
| | | | |
DB 667 CACGCCGCGCCCATCGCCCCCGGCGAGATCGCGAGCCCGCGGAGAGAGATCGCGGC 726
| | | | |
QY 709 ACCACAGCAGCCTCGAGAGAGAGATCGCTTGATGACCAAGAACCCCTCCATCCCGTG 768
| | | | |
DB 727 ACCACAGCAGCCTCGAGAGAGAGATCGCTTGATGACCAAGAACCCCTCCATCCCGTG 786
| | | | |

```

```

QY 769 GGGACATCTACAAAGGCGTGGATCATCTGGGGCTGAACAAGATCGTGGGATGTACAGC 828
DB 787 GGGAGATCTACAAAGCGTGGATCATCTGGGGCTGAACAAGATCGTGGGATGTACAGC 846
QY 829 CCCGTGAGATCTGTGACATCAAGAGGCGCCCAAGAGGCGCTTCCGCGCATCTAGTGAGC 888
DB 847 CCCACAGCATCTGTGACATCTGGGAGGCGCCCAAGAGGCGCTTCCGCGCATCTAGTGAGC 906
QY 889 CGCTTCTTCAAGACCTTGGCGCGCGAGCAGAGCAACCAGAGGTGAAGAACTGGATGACC 948
DB 907 CGCTTCTTCAAGACCTTGGCGCGCGAGCAGAGCAACCAGAGGTGAAGAACTGGATGACC 966
QY 949 GACACCTGTGTGTGAGAAAGCGCAACCCCGAATGTGAAGACCTCTGGCGCTCTGGC 1008
DB 967 GAAACCTGTGTGTGAGAAAGCGCAACCCCGAATGTGAAGACCTCTGGCGCTCTGGC 1026
QY 1009 CCCGCGCGCAGCTGTGAGAGATGATGACCGCTGCGAGGGCGTGGGCGGCGCCAGCAGC 1068
DB 1027 CCCGCGCGCAGCTGTGAGAGATGATGACCGCTGCGAGGGCGTGGGCGGCGCCAGCAGC 1086
QY 1069 AAGCGCGCGTGTGTGAGAGGCGGATGAGCAGGCGCAACA-----CAAGCTGATGATG 1122
DB 1087 AAGCGCGCGTGTGTGAGAGGCGGATGAGCAGGCGCAACA-----CAAGCTGATGATG 1146
QY 1123 CAGAAAGCAATCTTCAAGAGGCGCGCGCGCATGTCTCAAGTCTTCAACTGCGCAAGAG 1182
DB 1147 CAGCGCGCAACTCTTCCGCAACCGAGGAGACCGTCAAGTCTTCAACTGCGCAAGAG 1206
QY 1183 GGCACATGCGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242
DB 1207 GGCACATGCGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266
QY 1243 GAGGCGCACCAATGAGAGCTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302
DB 1267 GAGGCGCACCAATGAGAGCTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1326
QY 1303 CCCAGCACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1362
DB 1327 CCCAGCACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY 1363 CCCGCGCAGAGCTTCCGCTTC-----GAGAGAGCAACCCCGCGCGCGCGCGCGCGCG 1413
DB 1387 CCCGAGAGAGCTTCCGCTTC-----GAGAGAGCAACCCCGCGCGCGCGCGCGCGCG 1446
QY 1414 ---AGCAAGACCGCGAGACCTGACCAAGCTGAGAGGCTGTTCCGCGCAACGACCCCTG 1470
DB 1447 ATCGAACAGAGCTGTAACCCCTGACCAAGCTGCGAGCTGTTGGGCAACGACCCGACG 1506
QY 1471 AGCCAGT 1477
DB 1507 AGCCAGT 1513

```

# RESULT 7 US-09-475-515-5

```

; Sequence 5, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30

```

```

; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-09-475-515-5

Query Match      75.4%; Score 1114.6; DB 4; Length 1853;
Best Local Similarity 85.8%; Pred. No. 6,6e-164;
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 7 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 67 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 126
QY 121 CTGGAAGATTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 127 CTGGAAGATTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186
QY 181 ATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 187 CTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 246
QY 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 247 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY 301 CTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 307 CTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 366
QY 361 GCGG-----ACAAAGGCAAGGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
DB 367 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
QY 409 CAGGCGCGAGTGTGACCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
DB 427 CAGGCGCGAGTGTGACCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 469 ATCGAGAGAGAGGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 528
DB 487 GTGAGAGAGAGGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
QY 529 GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 588
DB 547 GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
QY 589 CAGATGCTGAAGAGACCAATCAACGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 648
DB 607 CAGATGCTGAAGAGACCAATCAACGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 666
QY 649 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 708
DB 667 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
QY 709 ACCACCGAGACCGCTGAGAGAGAGATGCGCTGATGACAGAGACCGCGCGCGCGCGCGCG 768
DB 727 ACCACCGAGACCGCTGAGAGAGAGATGCGCTGATGACAGAGACCGCGCGCGCGCGCGCG 786
QY 769 GCGGACATCTTCAAGAGGCGGTGATCATCTGGCGCTTGAACAAAGATGTGCGGATGACG 828
DB 787 GCGGACATCTTCAAGAGGCGGTGATCATCTGGCGCTTGAACAAAGATGTGCGGATGACG 846
QY 829 CCCGAGCATCTTGTGAGATCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888

```

Db 847 CCCACGACCTCTGACATCCGCCAGGCCCCAAGAGCCCTTCGCGACTACGTGAC 906  
Qy 889 CGCTTCTTAAGAACCTCTGCGCCGCGAGACCAAGAGAGTGAAGATGATGACC 948  
Db 907 CGCTTCTTAAGAACCTCTGCGCCGCGAGACCAAGAGAGTGAAGATGATGACC 966  
Qy 949 GACACCTGCTGTGAGAACCCCAACCCCGACTGCAAGACCACTCTGCGCTCTCGC 1008  
Db 967 GAGACCTGCTGTGAGAACCCCAACCCCGACTGCAAGACCACTCTGAGGCTCTCGC 1026  
Qy 1009 CCGCGCGCGAGCTGAGAGAGTGAATGATGACCTGCGCAAGGCGTGGGCGCCCGCCAC 1068  
Db 1027 CCGCGCGCGAGCTGAGAGAGTGAATGATGACCTGCGCAAGGCGTGGGCGCCCGCCAC 1086  
Qy 1069 AAGCCCGCGTGTGCGCGAGGCGATGACCGAGCCCAAC-----CGACGCTGATGATG 1122  
Db 1087 AAGCCCGCGTGTGCGCGAGGCGATGACCGAGTGAACCAACCCGCGCACTCATGATG 1146  
Qy 1123 CAGAAAGCACTTCAAGGCGCCCGCGCATGCTCAAGTGTCTCAACTGCGCAAGAG 1182  
Db 1147 CAGCGCGCACTTCCGCAACCGAGAGACGCTCAAGTGTCTCAACTGCGCAAGAG 1206  
Qy 1183 GCGCATGTGCGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242  
Db 1207 GCGCATGTGCGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266  
Qy 1243 GAGGCGCACTGAGTGAAGAGTGAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302  
Db 1267 GAGGCGCACTGAGTGAAGAGTGAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1326  
Qy 1303 CCGAGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1362  
Db 1327 CCGTCTTAAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386  
Qy 1363 CCGCGCGAGAGTCTCGCTTCAAGAG-----ACACCGCGCGCGCGCGCGCGCGCGCG 1413  
Db 1387 CCGAGAGAGAGTCTCGCTTCAAGAG-----ACACCGCGCGCGCGCGCGCGCGCGCG 1446  
Qy 1414 ---AGCAAGAGCGCGAGACCTGACCGAGCTGAAGAGCGCTGTGCGCAAGCGCGCTG 1470  
Db 1447 ATGAGCAAGAGAGTGTATCTTAACTTCTTCAAGTCACTCTTGTGCAAGCGCGCTG 1506  
Qy 1471 AGCCAGTAA 1479  
Db 1507 TCACAGTAA 1515

RESULT 8  
US-09-475-515-78  
; Sequence 78, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 78  
; LENGTH: 1865  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: GPI  
US-09-475-515-78  
Query Match 75.4%; Score 1114.6; DB 4; Length 1865;  
Best Local Similarity 85.8%; Pred. No. 6.6e-164;  
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;  
Qy 1 ATGGGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGAGCGCGCTGGAGCGCATCGCG 60  
Db 13 ATGGGCGCGCGCGAGCGTGTCTGAGCGCGCGCGCAAGCTGAGCAAGTGGAGAAATCGCG 72  
Qy 61 CTGCGCGCGCGCGCAAGAGTCTCAATGATGAAGCACTGTGTGGCCAGCCCGAG 120  
Db 73 CTGCGCGCGCGCGCAAGAGTCTCAATGATGAAGCACTGTGTGGCCAGCGCGAG 132  
Qy 121 CTGGAAGATTCGCGCGTGAACCGCGCGCTGTGAGACCAAGAGGCGTGAAGCATC 180  
Db 133 CTGGAAGATTCGCGCGTGAACCGCGCGCTGTGAGACCAAGAGGCGTGAAGCATC 192  
Qy 181 ATCCGCGAGTGAACCGCGCGCTGAGACCGCGCGAGAGAGTGAAGAGCTGTTCAC 240  
Db 193 CTGCGCGAGTGAAGCGCGAGCTGAGACCGCGCGAGAGAGTGAAGAGCTGTTCAC 252  
Qy 241 ACCGTGCGCGCGTGTACTGTGTCAGAGAAATCGAGTTCGCGGACCAAGAGGCG 300  
Db 253 ACCGTGCGCGCGTGTACTGTGTCAGAGAAATCGAGTTCGCGGACCAAGAGGCG 312  
Qy 301 CTGGAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Db 313 CTGGAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372  
Qy 361 GCGG-----ACAGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408  
Db 373 GCGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432  
Qy 409 CAGGCGAGATGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468  
Db 433 CAGGCGAGATGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492  
Qy 469 ATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528  
Db 493 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552  
Qy 529 GCGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588  
Db 553 GCGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612  
Qy 589 CAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648  
Db 613 CAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672  
Qy 649 CAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708  
Db 673 CAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732  
Qy 709 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768  
Db 733 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792  
Qy 769 GCGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828  
Db 793 GCGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852  
Qy 829 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888  
Db 853 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 912  
Qy 889 CGCTTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948  
Db 913 CGCTTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972  
Qy 949 GACACCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008

Db 973 GAGACCTCTGTGTGAGAAAGCCAAACCCGACCTGCAAGACATCTTGAAGGCTCTGGC 1032  
Qy 1009 CCGGCGCCGACCTGAGAGATGATGACCGCTTGCACAGGCGCTGGGCGGCCCAAGCCAC 1068  
Db 1033 CCGGCGCCGACCTGAGAGATGATGACCGCTTGCACAGGCGCTGGGCGGCCCAAGCCAC 1092  
Qy 1069 AAGGCGCGCTGTGCGGAGGAGGAGTGAAGCAGAGCCAAACA-----CCAGCTGATGATG 1122  
Db 1093 AAGGCGCGCTGTGCGGAGGAGGAGTGAAGCAGAGCCAAACA-----CCAGCTGATGATG 1152  
Qy 1123 CAGAGAGCACTTCAAGGAGCCCGCGCGATCGTCAAGTCTTCAACTGCGGCAAGAG 1182  
Db 1153 CAGCGCGCAACTTCCGCAACAGCGGAGACCGTCAAGTCTTCAACTGCGGCAAGAG 1212  
Qy 1183 GGGCAATGCGCCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242  
Db 1213 GGGCAATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1272  
Qy 1243 GAGGCGCACTGATGAGAGAGTGCACCGGAGCGCGCGCAACTTCTGGGCAAGATCTGG 1302  
Db 1273 GAGGCGCACTGATGAGAGAGTGCACCGGAGCGCGCGCAACTTCTGGGCAAGATCTGG 1332  
Qy 1303 CCGAGCGCAAGAGGCGCGCGCGCGCAACTTCTGCAAGCGCGCGCGCGCGCGCGCGCG 1362  
Db 1333 CTTCTTACAAAGGAGAGGCGCGGGAATTTTCTTCAAGGAGACCAAGGCGCAAGCGCCCA 1392  
Qy 1363 CCGGCGCAAGATCTTCCGCTTCAAGAG-----ACCACTCCGCGCGCAAGAGCGAG--- 1413  
Db 1393 CCAAGAGAGAGCTTCAAGGTTGGGAGAGAAACAATCTCTCAAGAGCGAGCGCG 1452  
Qy 1414 ---AGCAAGAGCGCGGAGACCTGACGAGCTGAGAGCGCTGTTGGGCAAGCGCGCGCG 1470  
Db 1453 ATAGACAAAGGACTGATCTTTAACTTCCCTCAAGTCACTTGTGGCAAGCGCGCTGG 1512  
Qy 1471 AGCCAGTAA 1479  
Db 1513 TCACAGTAA 1521

RESULT 9  
US-09-475-515-79  
; Sequence 79, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR WEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 79  
; LENGTH: 1865  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GP2  
US-09-475-515-79

Query Match 75.4%; Score 1114.6; DB 4; Length 1865;  
Best Local Similarity 85.8%; Pred. No. 6,6e-164;  
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

Qy 1 ATGGGCGCGCGCGCGCATCTGTGCGGCGGCAAGCTGAGCGCTGGAGGCGCATCCGC 60

Db 13 ATGGGCGCGCGCGCGCATCTGTGAGCGGCGGAGCTGAGCAAAAGTGGAGAAATCCGC 72  
Qy 61 CTGGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120  
Db 73 CTGGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCGCGAG 132  
Qy 121 CTGGAGAAATTTGGCCCTGAACCCCGGCTGTGTGAGACCAAGAGAGGCTGTCAAGCATC 180  
Db 133 CTGGAGCGCTTGGCGGTGAACCCCGGCTGTGTGAGACCAAGAGAGGCTGTCCGCAAGATC 192  
Qy 181 ATCCGCAAGCTGACACCGCGCGCTGAGACCGGCAAGAGAGCTGAGAGCGCTTCAAC 240  
Db 193 CTGGGCGCACTGACCGCGCGCTGAGACCGGCAAGAGAGCTGAGAGCGCTTCAAC 252  
Qy 241 ACCGTGGCCACCTGTACTGCTGTGACGAGAAATGAGTCCGCGACCAAGAGAGCGC 300  
Db 253 ACCGTGGCCACCTGTACTGCTGTGACGAGAAATGAGTCCGCGACCAAGAGAGCGC 312  
Qy 301 CTGGACAAATGTGAGAGAGAGCAAAATGTCACAGAGATTCAGAGGCGCGAGCGC 360  
Db 313 CTGGAGAAATGTGAGAGAGAGCAAAATGTCACAGAGAGGCGCGAGCGCGCGC 372  
Qy 361 GCGG-----ACAAAGGCGAAGGTGAGCGCAAACTACCCATGTGTGCAAACTGG 408  
Db 373 GCGGCGCGGCACTGGCAAGAGAGGAGTGAACACCGTGGGCGGCGCACAGGCGCGCATG 432  
Qy 409 CAGGCGCAGATGTGTACACAGGCGCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 468  
Db 433 CAGGCGCAGATGTGTACACAGGCGCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 492  
Qy 469 ATCGAGAGAGAGGCTTCAAGCGCGCGAGTATCCCATGTTTCAAGCGCGCGCGCGCG 528  
Db 493 GTGAGAGAGAGGCTTCAAGCGCGCGAGTATCCCATGTTTCAAGCGCGCGCGCGCG 552  
Qy 529 GCCACCGCGCGAGGAGCTGAAACAGATGTGAACACCGTGGGCGGCGCACAGGCGCGCATG 588  
Db 553 GCCACCGCGCGAGGAGCTGAAACAGATGTGAACACCGTGGGCGGCGCACAGGCGCGCATG 612  
Qy 589 CAGATGCTGAAGAGACCATCAACAGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCGCTG 648  
Db 613 CAGATGCTGAAGAGACCATCAACAGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCGCTG 672  
Qy 649 CAGCGCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 708  
Db 673 CAGCGCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 732  
Qy 709 ACCACAGAGACCGTGCAGAGAGAGTCCGCTGAGTGAACACGACACCGCGCGCGCGCG 768  
Db 733 ACCACAGAGACCGTGCAGAGAGAGTCCGCTGAGTGAACACGACACCGCGCGCGCGCG 792  
Qy 769 GCGGACATCTCAAGACCGGTGATCATCTGGGCGCTGAACAAGATGTGCGATGTAACG 828  
Db 793 GCGGAGATCTCAAGACCGGTGATCATCTGGGCGCTGAACAAGATGTGCGATGTAACG 852  
Qy 829 CCGGTGAGCATCTGGAACATCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888  
Db 853 CCGACAGACATCTGGAACATCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 912  
Qy 889 CGCTTCTTCAAGACCGTGCAGAGAGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948  
Db 913 CGCTTCTTCAAGACCGTGCAGAGAGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 972  
Qy 949 GACACCTGTGTGTGAGAGAGAGTGAACCGCGCTGCGACAGGCGTGGCGCGCGCGCG 1008  
Db 973 GAGACCTGTGTGTGAGAGAGAGTGAACCGCGCTGCGACAGGCGTGGCGCGCGCGCG 1032  
Qy 1009 CCGGCGCGCGCGCGCGAGAGAGTGAACCGCGCTGCGACAGGCGTGGCGCGCGCGCG 1068  
Db 1033 CCGGCGCGCGCGCGCGAGAGAGTGAACCGCGCTGCGACAGGCGTGGCGCGCGCGCG 1092  
Qy 1069 AAGGCGCGCGTGTGGCGAGAGAGTGAACCGCGCTGCGACAGGCGTGGCGCGCGCG 1122



Db 1093 AAGCCCGCTGCGCCGAGCGATGAGCCAGTGAACGACC CGGCGACATCATGATG 1152  
Qy 1123 CAGAAAGAACTTCAAGGGCCCCCGCCCATCTCTCAAGTGTCTCAACTGCGCAAGAG 1182  
Db 1153 CAGCGCGGAATCTTCGCAACGAGCGGAAGACCGTCAAGTGTCTCAACTGCGCAAGAG 1212  
Qy 1183 GGCACATGCGCCGCAACTGCGCGCGCCCCCGCAAGAGAGGCTGTGAAATGCGGCAAG 1242  
Db 1213 GGCACACCGCGAGAACTGCGCGCGCCCCCGCAAGAGAGGCTGTGCGCGCGCGC 1272  
Qy 1243 GAGGCGACCATGATGAAGATGTCACCGAGCGCGCAACTTCTGCGCAAGATCTG 1302  
Db 1273 GAAAGCACCAATGAATGAAATGTCATGAGAGCAGGCTAATTTTAAAGGAAGATCTG 1332  
Qy 1303 CCAGCGCAAGAGGCGCGCGCGCACTTCTGCAAGCGCGCGGCGCGCGCGCGCGCG 1362  
Db 1333 CTTTCTCAAGGAAGAGCGCGAGAAATTTTCTTCAAGAGACGAGCGCAAGCGCGCA 1392  
Qy 1363 CCGCGGAGAGCTTCGCTTCGAGAG-----ACCACCGCGCGCGCAAGAGAGAG-- 1413  
Db 1393 CAGAAAGAGACTTCAGTGTGGGAGAGAAACAACTCCTCTCAAGAGAGAGAGCGG 1452  
Qy 1414 ---AGCAAGACCGCGAGACCTGACCAAGCTGAAGAGCTGTTCGCAAGACCGCTTG 1470  
Db 1453 ATAGACAAAGAACTGTATCTTTAACTTCTCAGATCACTCTTGGCAAGACCGCTCG 1512  
Qy 1471 AGCCAGTAA 1479  
Db 1513 TCACAGTAA 1521

RESULT 10  
US-09-475-515-6  
Sequence 6, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 4319  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: HIV-Gag-polymerase  
US-09-475-515-6

Query Match 75.4%; Score 1114.6; DB 4; Length 4319;  
Best Local Similarity 85.8%; Pred. No. 6,6e-164;  
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;  
Qy 1 ATGGGCGCCGCGCGCAGCATCTGCGCGGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
Db 7 ATGGGCGCCGCGCGCAGCGTGTGAGCGGCGGAGCTGAGCAAGTGGAGAAATCCGC 66  
Qy 61 CTGGCGCCGCGCGCGCAAGAGTGTCTACATGATGAAGCACCTGTGTGGCCAGCCGCGAG 120  
Db 67 CTGGCGCCGCGCGCGCAAGAGTGTCAAGCTGAACACATGTGTGGCCAGCCGCGAG 126

Qy 121 CTGGAAGATTGCGCTTGAAACCCCGGCTCTGTGAGACACAGCGAGGCTGCAAGCAGATC 180  
Db 127 CTGAGAGCGCTTGCGGTGAACCCCGGCTCTGTGAGACACAGCGAGGCTGCGCAGATTC 186  
Qy 181 ATCCGCACTGTCACCCCGCTTGACACCGGCAAGCGGAGGAGCTGMAAGCTTTCAAC 240  
Db 187 CTGGGCGAGCTGAGCCCGAGCTGACACCGGCGAGGAGGAGCTGCGCAGCTGTACAAAC 246  
Qy 241 ACCGTGCGCACCGTGTATGTGCTGACAGAAAGATTCGAGTCCGCGACACCAAGAGGCG 300  
Db 247 ACCGTGCGCACCGTGTATGTGCTGACACCGCATTCAGCTCAAGAGACCAAGAGGCGC 306  
Qy 301 CTGCAAAAGATCGAGAGAGACAAACAAAGTCCAGCAGAAAGATTCAGCAGCGAGCGC 360  
Db 307 CTGGAAGAGATCGAGAGAGACAAACAAAGTCCAAAGAAAGAGGCGCAGCAGCGCGCGC 366  
Qy 361 GCGG-----ACAGGGCAAGGTGAGCCGAACCTACCCCATGTGTGCAAACTTG 408  
Db 367 GCGGCGGCGACCGGCAACAGAGCGCAGGTGAGCCAGAACTACCCCATGTGTGCAAACTTG 426  
Qy 409 CAGGCGCAGATGTGTCACAGGCGCATGACCCCGCGCACCTGAAAGGCTGAGTGAAGTG 468  
Db 427 CAGGCGCAGATGTGTCACAGGCGCATGACCCCGCGCACCTGAAAGGCTGAGTGAAGTG 486  
Qy 469 ATCGAGAGAAAGGCTTCAGCCCGAGGTGATCCCATGTTCACCGCTTGAGCGAGGCG 528  
Db 487 GTGAGAGAAAGGCTTCAGCCCGAGGTGATCCCATGTTCAGCGCTTGAGAGAGGCG 546  
Qy 529 GCGACCCCGCGAGACTTGAAACGATGTTGAACACCGTGGCGGCGCACAGCGCGCATG 588  
Db 547 GCGACCCCGCGAGACTTGAAACGATGTTGAACACCGTGGCGGCGCACAGCGCGCATG 606  
Qy 589 CAGATGTGAAGAGACACCATCAAGAGAGGCGCGGAGTGGGACCGGTGACACCGCTG 648  
Db 607 CAGATGTGAAGAGACCATCAAGAGAGGCGCGGAGTGGGACCGGTGACACCGCTG 666  
Qy 649 CAGCGCGGCGCCATGCGCCCGCGCAGATGCGGAGCGCCCGCGCAGCGACATGCGCGCG 708  
Db 667 CAGCGCGGCGCCATGCGCCCGCGCAGATGCGGAGCGCCCGCGCAGCGACATGCGCGCG 726  
Qy 709 ACCACAGACCTCTGAGAGAGCAATGCGCTGATGACAGCAACCCCGCATCCCGCTG 768  
Db 727 ACCACAGACCTCTGAGAGAGCAATGCGCTGATGACCAACCCCGCATCCCGCTG 786  
Qy 769 GCGGCACTTCAAGCGGTGATCATCTGTGGCGCTGAAACAAGATCTGCGGATGACAGC 828  
Db 787 GCGGATCTTCAAGCGGTGATCATCTGTGGCGCTGAAACAAGATCTGCGGATGACAGC 846  
Qy 829 CCGGTGAGCATCTCTGAGCATCAAGCAGGCGCCCAAGAGGCGCTTCGCGCATAGCTGAGC 888  
Db 847 CCCACAGCATCTCTGAGCATTCGCGCAGGCGCCCAAGAGGCGCTTCGCGCATAGCTGAGC 906  
Qy 889 CGCTTCTTCAAGACCTGCGCGCGCGAGAGCAACCGAGAGTGAAGAACTGATGAC 948  
Db 907 CGCTTCTTCAAGACCTGCGCGCGCGAGAGCAACCGAGAGTGAAGAACTGATGAC 966  
Qy 949 GACACCTCTGCTGTGACAAAGCGCAACCCCGCATCTGCAAGACCATCTCGGCGCTCTGCGC 1008  
Db 967 GAGACCTCTGCTGTGACAAAGCGCAACCCCGCATCTGCAAGACCATCTCGGCGCTCTGCGC 1026  
Qy 1009 CCGGCGCGCAGCTGTGAGAGATGATGACCGCTGCGCAGGCGGTGGGCGCGCCAGCCAC 1068  
Db 1027 CCGGCGCGCAGCTGTGAGAGATGATGACCGCTGCGCAGGCGGTGGGCGCGCCAGCCAC 1086  
Qy 1069 AAGGCGCGCTGTGCGCGAGGCGATGAGCCAGGCGCA-----CCAGCGTGAATG 1122  
Db 1087 AAGGCGCGCTGTGCGCGAGGCGATGAGCCAGGAGAGAACCGGCGCAGCATATGATG 1146  
Qy 1123 CAGAAAGCAACTTAAGGCGCCCGCGCATCTGCAAGTGTTCATCTGCGCGCAAGAG 1182  
Db 1147 CAGCGCGCGCACTTCTCGCAACAGCGGAAAGACCGTCAAGTGTTCATCTGCGCGCAAGAG 1206  
Qy 1183 GGCACATGTGCGCGCAACTGCGCGGCGCCCGCGCAAGAGGCTGTGGAAGTGGCGCAG 1242



Db	1207	GGCCACACCGCAGAACTGCCCGCCCCCGCAAGAGGGCTCTGGCGCTGGCGCGC	1266
Qy	1243	GAGGGCCACCATGATGAAGACTGCACCGAGCGCAGGCACTTCTGGGCAAGATCTGG	1302
Db	1267	GAGGACACCCAAATGMAAGATTGCATGAGAGACAGGCTAAATTTTTRAGGAAATCTGG	1326
Qy	1303	CCGAGCCACAAAGGCGCGCCCGCAACTTCTGAGAGCGCGCCCGAGCCACCGCCC	1362
Db	1327	CTTCTCTTCAAGGGAGAGGCGCAGGGAATTTTCTTCAAGACAGACAGCCCAAGCCCA	1386
Qy	1363	CCGCGCCAGAGACTTCCGCTTCGAGAG-----ACCAACCCCGCGCCAGAACAGAGAG---	1413
Db	1387	CGAGAGAGACTTTCAGGTTTGGGAGGAGAAACAATCTCTCTCAGAGACAGAGACCG	1446
Qy	1414	---AGCAAGACCGCGAGACCCCTGACCAAGCTGGAAGACTGTTCGGCAACGACCCCTG	1470
Db	1447	ATAGACAAAGAACTGTATCTTTAACTTCCCTCAGATCACTCTTTGGGAAGACCCCTCG	1506
Qy	1471	AGCCAGTAA	1479
Db	1507	TCACAGTAA	1515

## RESULT 11

```

US-09-475-515--9
; Sequence 9, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: common region
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; US-09-475-515--9

```

Query Match 69.38; Score 1025.2; DB 4; Length 1268;

	Matches 1131;	Conservative	0;	Mismatches 113;	Indels	18;	Gaps	2
Qy	1	ATGGGCGCCCGCGCCAGATCTCTGGCGCGCGCGAAGCTTGAGCGCTTGAGGAGCGCATCCCG	60					
Db	7	ATGGGCGCGCCCGCGCCAGCGTGCTGTGAGCGGCGCGCGCTTGACAAAGTGGGAGAAAGATCCCG	66					
Qy	61	CTGGGCGCCCGCGCGCAAGATGCTACATGATGAAGGCATCTGTGTGGGCGCAGCCGCGAG	120					
Db	67	CTGGCGCCCGCGCGGAGAGAAAGTACAAAGCTGAAGGCATATGTTGGGCGCAGCGCGAG	126					
Qy	121	CTGGAGAAAGTTGCGCCCTGAACCCCGGCGTGTGAGAGACCAAGCGAGGGCTGCAAGGATC	180					
Db	127	CTGGAGCGCCTTGCGCTGTAACCCCGGCTGTGTGAGACCAAGGAGGGCTGTCCGCGAAGATC	186					
Qy	181	ATCCGCGACGTGCACCCCGCCCTTGACAGCCGGCAGCGAGAGCTGAAGAGCTGTTCAAC	240					
Db	187	CTGGGCGCAGCTGCAGCCCGAGCTGTGAGACCGGCAACGAGAGACTGTGCGCAGCTGTGAACAAC	246					

QY	241	ACCGTGGCCACCTGTACTGTGTGACAGAAAGATTCGAGGTTCGGACACACCAAGAGGCC	300
Db	247	ACCGTGGCCACCTGTACTGTGTGACAGAGGCATTCGATCGATCAAGAACCAAGAGGCC	306
QY	301	CTGCAACAAGATTCGAGAGGAGACAGAACAAAGTGCAGACGAAGATCAGACAGGCC	360
Db	307	CTGAGAAAGATTCGAGAGGAGACAGAACAAAGTTCGAAGAAAGGCCACAGGCC	366
QY	361	GCCG-----ACAGGGCAGGTGAGGCCAGAACTAACCCATGTGTGCAACTTG	408
Db	367	GCCGCGGACACCGGCAACAGCAGCGCGGTGAGCCAGAACTAACCCATGTGTGCAACTTG	426
QY	409	CAGGGCCAGATGTGTGACAGGACCATCAGGCCCGGACCCCTGAAGAGCCGTGGGTGAAGTG	468
Db	427	CAGGGCCAGATGTGTGACCAAGGCATCAGCCCCCGGACCTTGAAACGCTTGGGTGAAGTG	486
QY	469	ATCGAGAGAAAGGCTTTCAGGCCCGGAGGTGATCCCATGTTCAACGCCCTTGAGCGAGGC	528
Db	487	GTGAGAGAGAAAGGCTTTCAGCCCCGAGGTGATCCCATGTTCAAGGCGCTTGAGCGAGGC	546
QY	529	GCCACCCCTCCAGAGACTTGAACAACCATGTTTGAACACCTGTGGCCGCTACACAGGCCCATG	588
Db	547	GCCACCCCTCCAGAGACTTGAACAACCATGTTTGAACACCTGTGGCCGCTACACAGGCCCATG	606
QY	589	CAGATGTGGAAGACACCATCAACGAGAGGCGCGCGATGAGGACCGGTGCAACCCCGTG	648
Db	607	CAGATGTGGAAGAGACCATCAACGAGAGGCGCGCGATGAGGACCGGTGCAACCCCGTG	666
QY	649	CACGCGGCGCCCATGCGCCCGCGCCAGATGCGGAGGCCCGCGGACGACATCGCGGC	708
Db	667	CACGCGGCGCCCATGCGCCCGCGCGAGATGCGGAGGCCCGCGGACGACATCGCGGC	726
QY	709	ACCAACACACACCTTCGAGAGACAGATGCGCTGATGACACACAAACCCCCATCCCGCTG	768
Db	727	ACCAACACACACCTTCGAGAGACAGATGCGCTGATGACACACAAACCCCCATCCCGCTG	786
QY	769	GCGCAGATCTTCAAGCGGTGATCATCTGGGCGCTGAACAAGATCTGTGCGATGTACAGC	828
Db	787	GCGCAGATCTTCAAGCGGTGATCATCTGGGCGCTGAACAAGATCTGTGCGATGTACAGC	846
QY	829	CCCGTAGCATCTTCGACATCAAGCAGGGGCCCAAGAGGCCCTTTCGCGACTACGTGCAC	888
Db	847	CCCAACACACATCTTCGACATCTTCGCGCAGGGGCCCAAGAGGCCCTTTCGCGACTACGTGCAC	906
QY	889	CGCTTCTTCAAGACCCCTGCGCGCGAGAGAGACCCGAGAGGTGAACATCGATGAC	948
Db	907	CGCTTCTTCAAGACCCCTGCGCGCTGAGAGAGCCGACCGAGACGTGAAGATCTGATGACC	966
QY	949	GACACCTGTGTGTGACAGAACGCCAACCCCGACTGCAAGACCATCTGTGCGCTTCGAGC	1008
Db	967	GAGACCTGTGTGTGACAGAACGCCAACCCCGACTGCAAGACCATCTGTGAGGCTTCGAGC	1026
QY	1009	CCCGCGCGCAACCTTCGAGAGATATATGACCGCTCCGACGAGCGCGTGGCGGCCACGAC	1068
Db	1027	CCCGCGCGCAACCTTCGAGAGATATATGACCGCTCCGACGAGCGCGTGGCGGCCACGAC	1086
QY	1069	AAGGCGCGCTGTGCGCGAGGCGATGAGCGAGGCCAAC-----CCAGCGTATGATG	1122
Db	1087	AAGGCGCGCGTGTGTGCGAGGCGATGAGCGAGGCGAACCCGCGGACCATATGATG	1144
QY	1123	CAGAAAGCAACTTTCAMAGGGCCCCCGCGCATCTGCAATGCTTTCATCTGCGCAAGAG	1182
Db	1147	CAGGCGCGCAACTTTCGCAACCAACGAGAGACCGTCAAGTCTTTCATCTGCGCAAGAG	1206
QY	1183	GAGCCACATCGCGCGCAACTGCGCGGCCCGCGCAAGAAAGGCGCTGTGGAATGTCGCGCAAG	1242
Db	1207	GAGCCACATCGCGCGCAACTGCGCGGCCCGCGCAAGAAAGGCGCTGTGCGCGCGCGC	1266
QY	1243	GA 1244	
Db	1267	GA 1268	

```
RESULT 12
US-09-952-060-27
Sequence 27, Application US/09952060
Patent No. 6733993
GENERAL INFORMATION:
APPLICANT: Emilio A.
APPLICANT: Youli, Rima
APPLICANT: Betz, Andrew J.
APPLICANT: Chen, Ling
APPLICANT: Kaslow, David C.
APPLICANT: Shiver, John W.
APPLICANT: Toner, Timothy J.
APPLICANT: Casimiro, Danilo R.
TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV-1 GAG, POL, NEF AND
TITLE OF INVENTION: MODIFICATIONS
FILE REFERENCE: 20747Y
CURRENT APPLICATION NUMBER: US/09/952,060
CURRENT FILING DATE: 2001-09-14
PRIOR APPLICATION NUMBER: PCT/US01/28861
PRIOR FILING DATE: 2001-09-14
PRIOR APPLICATION NUMBER: 60/317,814
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 60/279,056
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 60/233,180
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 1521
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Codon optimized DNA encoding human HIV-1 gag
US-09-952-060-27

Query Match      66.0%; Score 976.6; DB 4; Length 1521;
Best Local Similarity 80.2%; Pred. No. 1.4e-142;
Matches 1205; Conservative 0; Mismatches 274; Indels 24; Gaps 4;

QY      1 ATGGGCGCCGCGCGCAGATCTCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
DB      1 ATGGGCTAGAGGCTTCTGTGCTGTGTGTGAGTGAAGCAAGTGGAGAAATCAAG 60
QY      61 CTGGCCCCCGCGCGCAAGAGTCTACATGATGAAACACTGTGTGTGGCCAGCGCGAG 120
DB      61 CTGAGGCTGTGTGGCAAGAAAGTACAAAGCTAAAGCAATTGTGTGGGCTTCCAGGGAG 120
QY      121 CTGGAAGATTGCGCCTGAAACCCCGGCTGTGTGAGAGACGAGGAGGCTGCAAGCATTC 180
DB      121 CTGGAGAGTTTGTCTGTGAACCTTGGCTGTGTGAGAGACTTGTAGGGGTGCGAGCATTC 180
QY      181 ATCCGCGACCTGACCCCGCCTCTGACAGACCGGCGAGGAGCTGAAAGCTGTTCAC 240
DB      181 CTGGGCGACCTCGAGCCTCTCCCTGCAACAGGCTGTGAGAGCTGAGTGTCTCCCTGTACAC 240
QY      241 ACCGTGGCAACCTGTACTGTCTGTGACAGAAAGATGAGTCCGCGACACCAAGAGGCTC 300
DB      241 ACGATGTGTAACCTGTACTGTGTGTGACCAAGAAATTGATGTAAAGACCAAGAGAGGCTC 300
QY      301 CTGACCAAGATCGAGGAGGAGCAAGCAAGTGTGACAGCAAGATCCAGGAGGCGCGAGG-- 358
DB      301 CTGGAGAGATTGAGAGAGAGCAAGCAAGTCCAGAGAGAGGCCAGAGGCTGTGCTT 360
QY      359 ----CGCGCAGCAAGGCAAGGTGAGCCAGAACTACCCCATGTGTGCAAGACCTGCAAGGC 414
DB      361 GGCACAGGCACTCCAGCCAGGTGTCCAGAACTACCCCATGTGTGCAAGACCTCCAGGCTC 420
QY      415 CAGATGTGTGACCAAGGCTCATCAAGCCCGGCACTCTGAAGCTGTGGTGAAGTGTATCGAG 474
DB      421 CAGATGTGTGACCAAGGCTCATCTCCCGGCACTCTGAAGTGTGGTGAAGTGTGTGAG 480
```

```
QY      475 GAGAGGCTTTCAGGCCCCGAGGTGATCCCAATGTTCAACGCGCTGAGGAGGCGCCAC 534
DB      481 GAGAGGCTTTCCTCCCTGAGGTGATCCCAATGTTCTTGTGCTGTCTGAGGAGTGCACC 540
QY      535 CCCCAGGACTTGAACACAGATTTGAACACCGTGGCGGGCCACAGGCGCCCATGACATG 594
DB      541 CCCCAGGACTTGAACACAGATCTGAACACAGTGGGGGGGCATCAAGGCTGCCATGACATG 600
QY      595 CTGAAGGACACCATTAAGAGAGAGCGCCGAGTGTGGACCGCTGTGACCCCGTGCACGCC 654
DB      601 CTGAAGGAGACCATTAAGAGAGAGCTGTGATGTGGACAGGCTGTGACCTGTGTGACCT 660
QY      655 GGGCCCATTCGCCCCCGGCGAGATGCGAGCGCCGCGCAGCAGCATTCGCGGCGCAC 714
DB      661 GGGCCCATTCGCCCCCGGCGAGATGAGGAGCGCCGAGGCTGTGACATTCGCGGCGCAC 720
QY      715 AGCACCTTGCAGAGACAGATGCGCTGTGATGACACAGCAACCCCTCATCCCGTGGCGAC 774
DB      721 TCACCCCTTCAGAGACAGATTTGGCTGTGATGACCAACACCCCTCATCCGTGGGGGAA 780
QY      775 ATCTACAGCGGTGATGATCTGGGCTGTGAACAAAGATCGTGGGATGTACAGCCCTG 834
DB      781 ATCTACAGAGGTGATGATCTGGGCTGTGAACAAAGATTTGAGAGATGTACTCCCCAC 840
QY      835 AGCATCTGTGACATTCAGACAGGCGCCCAAGAGCGCTTCCGCGACTACGTGAGACGCTTC 894
DB      841 TCATCTGTGACATTCAGACAGGCGCCCAAGAGCGCTTCCAGGACTATGTGACAGCTTC 900
QY      895 TTCAAGACCTTGGCGCGCGAGCAAGACCCAGAGAGGTGAAGAACTGTGATACCGACAC 954
DB      901 TACAAGACCTTGAAGGCTGAGACAGGCGCTCCAGAGAGGTGAAGAACTGTGATACAGAGAC 960
QY      955 CTGCTGTGTGAGAAACCGCAACCCGACTGTGAAGCAATCTCTGCGCGCTGTGGCCCGGC 1014
DB      961 CTGCTGTGTGAGAAATGCAACCTGTGACTGAAGCAATCTGTGAAGCGCTGTGGCCCTGCT 1020
QY      1015 GCCAGCTGTGAGAGATGATGACCGCTGTGCAAGGCGGTGGGCGGCCCGCAGCAAGGCGC 1074
DB      1021 GCCAGCTGTGAGAGATGATGACAGCTGTGCAAGGCGGTGGGCGGCCCTGTGTCAAGGCGC 1080
QY      1075 CGCGTGTGCGCGAGGCGATGAGCGCAAGCA-----CGACGTGTGATGTGCAAG 1128
DB      1081 AGGGGTGTGCGCTGAGGCGCATGTCCAGGTGACCAATCCGCAACATCATGATGTGCAAG 1140
QY      1129 AGCAATCTCAAGGCGCCCGGCGCATGTGCAAGTGTCTCAACTGTGCGCAAGAGGCGCAC 1188
DB      1141 GGCAACTTCAGGAAACAGAGGAAAGCAAGTGAAGTCTTCAACTGTGTGCAAGGTGGGCGAC 1200
QY      1189 ATCGCCCGCAACTGCGCGCGCCCGCGCAAGAAAGGCTGTGGAAGTGTGCGCAAGAGGCGC 1248
DB      1201 ATTGCGCAAGAACTGTAGAGGCGCCCGAGAGAAAGGCTGTGGAAGTGTGCGCAAGAGGCGC 1260
QY      1249 CACGATGAAGAGACTGACCGAGCGCGCAACTTCTGTGGCAAGATTTGCGCCAC 1308
DB      1261 CACGATGAAGAGACTGCAATGAGAGGCGAGGCAACTTCTGTGGCAAGATTTGCGCCCTCC 1320
QY      1309 CACAAGGCGCGCCCGGCAACTTCTGTGAGAGCGCCCGGCGCAACCGCCCGCCCGGCC 1368
DB      1321 CACAAGGCGCGCGCTGCGCAACTTCTGTGAGAGCGCTGAGGCCCAAGCGCCCTGCCAG 1380
QY      1369 GAGAGCTTTCGCTT-----CGAGAGACCAACCCCGCGCAGAGCAGAG-----AGC 1416
DB      1381 GAGTCTTTCAGGTTTGGGAGAGAGAGACCAACCCCGCGCAGAGAGCCCATTTGAC 1440
QY      1417 AAGAGCGCGAGACCTTGACCAAGCTGTGAAGCTGTTCGGCAACGACCCCTTGAGCCAG 1476
DB      1441 AAGAGCTGTACCCCTGTGCTCTGTGTGAGTCCCTGTGTGCAAGACCCCTCTCCAG 1500
QY      1477 TAA 1479
DB      1501 TAA 1503
```



Db 1270 CACGAGATGAAGACTGCAATGAGAGGAGCCCACTTCTGGGCAAAATCTGGCCCTCC 1329  
Qy 1309 CACAAGGGGCCCCCGGCAACTTCTGCAAGAGCCGCCGCAACCCGCCCCCGGCC 1368  
Db 1330 CACAAGGGGAGGCTGGCAACTTCTGCAAGAGCCGCCGCAACCCGCCCCCGGCC 1389  
Qy 1369 GAGAGCTTCCGCTT-----CGAGAGACCAACCCCGGCGCAAGAGAGAG-----AGC 1416  
Db 1390 GAGCTCTTAGGTTGGGAGAGAGACCAACCCGCAAGAGAGAGCCCATTCAGC 1449  
Qy 1417 AAGGACCGGAGACCTGACCAAGCTGAAGAGCTGTTGGGACGACCCCTGAGCCAG 1476  
Db 1450 AAGGAGCTTACCCCTGGGCTTCCCTGAGGTCCCTGTTGGCAACGACCCCTCTCCAG 1509  
Qy 1477 TAA 1479  
Db 1510 TAA 1512

RESULT 14  
US-09-818-443-1  
Sequence 1, Application US/09818443  
Patent No. 6787351  
GENERAL INFORMATION:  
APPLICANT: Chen, Ling  
APPLICANT: Shiver, John W.  
APPLICANT: Betz, Andrew J.  
APPLICANT: Casimiro, Danilo R.  
APPLICANT: Caulfield, Michael J.  
APPLICANT: Chaestain, Michael A.  
APPLICANT: Emili, Emilio A.  
TITLE OF INVENTION: ADENOVIRUS CARRYING GAG GENE HIV VACCINE  
FILE REFERENCE: 20440Y1A  
CURRENT APPLICATION NUMBER: US/09/818,443  
CURRENT FILING DATE: 2001-09-13  
PRIOR APPLICATION NUMBER: PCT/US00/18332  
PRIOR FILING DATE: 2000-07-03  
PRIOR APPLICATION NUMBER: 60/148,981  
PRIOR FILING DATE: 1999-08-13  
PRIOR APPLICATION NUMBER: 60/142,631  
PRIOR FILING DATE: 1999-07-06  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1532  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Optimized human HIV-1 gag ORF  
US-09-818-443-1

Query Match 66.0%; Score 976.6; DB 4; Length 1532;  
Best Local Similarity 80.2%; Pred. No. 1,4e-142;  
Matches 1205; Conservative 0; Mismatches 274; Indels 24; Gaps 4;

Qy 1 ATGGGGCGCGCGCGCGAGATCTGCGCGCGCGCAAGCTGAGCGCTGGAGGCGATCCGC 60  
Db 10 ATGGGCTGTAAGGCTTCTGTGCTGTGTGGTGGAGCTGGAACAAGTGGAGAAATCAGG 69  
Qy 61 CTGGCGCGCGCGCGCAAGAGTGTCTAATGATGAAAGCACTGGTGTGGCCAGCCGCGAG 120  
Db 70 CTGAGGCTGTGTGGCAAGAAAGTACAGCTTAAGCACTATTGTGGCTCTCCAGGAG 129  
Qy 121 CTGAGAAATTTGGCTCTGAAACCCCGGCTGTGTGAGACCAAGAGGGCTGCAAGCATC 180  
Db 130 CTGGAGAGTTTGTCTGTGAACCTTGGCTGTGTGAAACCTTGTAGGGGTGCAAGCATC 189  
Qy 181 ATCCGCGAGCTGACCCCGGCTGTGAGACCGGCGAGGAGAGCTGAAGGCTGTTCAGC 240  
Db 190 CTGGGCGAAGCTCAAGCTCTCTGCAAAAGGCTCTGAGAGAGCTGAGTCCCTGTACAC 249  
Qy 241 ACCGTGGCAACCTGTACTGCGTGCAGAAAGATGAGGTCCGCGACCAAGAGGCG 300

Db 250 AAGTGGCTACCTGTACTGTGTGCAACAGAAAGATTGATGTGAAGACCAACGAGAGGCC 309  
Qy 301 CTGGCAAGATTCGAGAGAGAGACCAAGAAATGTCACAGCAAGAAATTCACAGAGCCGAGG-- 358  
Db 310 CTGGAGAAATTTGAGAGAGAGACCAAGAAATGTCAGAAAGAAAGGCGCCAGACAGCTGTGCT 369  
Qy 359 -----CCGCGCAAGAGGCGAAGGTGAGCGAAGTAACTAACCCATCTGTGCAAGAACTTCGAGGC 414  
Db 370 GGCAAGGCAACTCCAGCGCAAGTGTCCAGAACTAACCCATTTGTGCAAGAACTTCAGAGGC 429  
Qy 415 CAGATGTGCAACAGGCGCATAGCCCCCGCAACCTGAAAGCTTGGGTGAAGGTATGAG 474  
Db 430 CAGATGTGCAACAGGCGCATCTCCCCCGGACCTGAATGCTGGGTGAAGGTATGAG 489  
Qy 475 GAGAAAGGCTTCAGGCGCGAAGGTATCCCAATGTTACCGGCTGAGGAGAGGGCGCAC 534  
Db 490 GAGAAAGGCTTCCTCCCTGAGGTATCCCAATGTTCTGTGCTGTGAGAGGTGCGCAC 549  
Qy 535 CCGCAGACCTGAACACAGATGTTGAACACGCTGGGCGGCGCACAGGCGCGCATGAGATG 594  
Db 550 CCGCAGACCTGAACACAGATGTTGAACACAGTGGGGGGCGCATCAGGCTGCCATGCAATG 609  
Qy 595 CTGAAGGACACCATCAACGAGAGGCGCGGAGTGGAGACCGGTGACACCCGCTGACGCC 654  
Db 610 CTGAAGGAGACCATCAATGAGAGGCTGTGAGTGGAGACAGGCTGATCTGTGACGCT 669  
Qy 655 GGGCCCATTCGCCCCCGGCAATGCGGAGCCCGCGGCGACGACATGCGCGGACCAACC 714  
Db 670 GGGCCCATTCGCCCCCGGCAATGAGGAGGCGCGAGGCTGTGACATTTGTGCGCACCAACC 729  
Qy 715 AGCACCTGCGAGAGACAGATGCGCTGATGACCAAGCAACCCCGCATCCCGGTGGGCGAC 774  
Db 730 TCACACCTCGAGAGACATTTGGCTGATGACCAACACCCCGCATCTCTGTGGGGGAA 789  
Qy 775 ATCTAAGACGGGTGATATCTTGGGCTGAAACAAGATGTCGAGATGACGCCCGTG 834  
Db 790 ATCTAAGAGGTGATATCTTGGGCTGAAACAAGATGTCGAGATGATCTCCCGCAC 849  
Qy 835 AGCATCTGAGACATCAACAGAGGCGCCCAAGAGGCTTCCGCGACTACGTGAGACCGCTTC 894  
Db 850 TCACATCTGAGACATCAACAGAGGCGCCCAAGAGGCTTCAAGGACTAATGTGAGACAGGTT 909  
Qy 895 TTCAAGACCTTGGCGCGCGAGACAGACCAAGAGGTGAAGAACTGGATGACGACAC 954  
Db 910 TACAAAGACCTTGAAGGCTGAGACAGGCTCTCCAGAGAGTGAAGAACTGGATGACAGAGAC 969  
Qy 955 CTGCTGTGCAAGAGCGCAACCCCGCATGCAAGACCATCTTGGCGGCTCTGCGCCCGGC 1014  
Db 970 CTGCTGTGCAAGATGCAACCTGACCTGCAAGACCATCTGAAAGGCTCTGCGGCTGTGCT 1029  
Qy 1015 GCCAGCCTGAGAGATGATGACCGGCTGCAAGGCGGTGGGCGGCCCAAGCACAAGGCC 1074  
Db 1030 GCCAGCCTGAGAGATGATGACAGCTGCAAGGAGGCGGTGGGCGGCCCTGTGTCACAAGGCC 1089  
Qy 1075 CGCGTGTGCGCGAGGCGATGAGCGAGCGCAACA-----CCAGCTGATGATGACAGAG 1128  
Db 1090 AGGCTGTGCTGAGAGCATGTCAGAGTACCAAGTACCAACTCCGCAACATATGATGACAGAG 1149  
Qy 1129 AGCAACTTCAAGAGGCGCGCGCATGCTCAAGTGTCTCAACTCGCGGCAAGAGGCGCAC 1188  
Db 1150 GGCAACTTCAAGAACTCAAGAGAGACAGTGAAGTGTCTCAACTGTGGCAAGTGGGCGCAC 1209  
Qy 1189 ATCGGCGCGCATGCGCGGCGCGCGCAAGAGGCGCTGGAAGTGGGCGCAAGAGGCG 1248  
Db 1210 ATTGCAAGAACTGAGAGGCGCGCGCAAGAGGCGCTGGAAGTGGGCGCAAGAGGCG 1269  
Qy 1249 CACGAGTGAAGAGCTGACCGAGGCGCAAGCTTCTGTGGCAAGATGTCGCCAGC 1308  
Db 1270 CACGAGTGAAGAGCTGCAATGAGAGGCGCAACTTCTGTGGCAAAATCTGTGCGCTCC 1329  
Qy 1309 CACAAGGCGCGCGCGCAACTTCTGTGAGAGCGCGCGGAGCCACCGCCCGCC 1368  
Db 1330 CACAAGGCGAGGCTGGCAACTTCTGTGAGAGCGCGCTGAGCCACAGCCCTTCCGAG 1389

```

Qy      1369 GAGACCTTCGCGTT-----CGAGAGCAACCCCCGGCCGAAGCAGAG------AGC   1416
Db      1390 GAGTCTCTTACAGATTGGGAGAGAAAGCACCCCACGCCCTCGAAGCAGAGCCCATTTGAC   1449
Qy      1417 AAGGACCGCGAGACCCTTGACAGGCCTGTGAAGAAGCCGTTCGGCAACGACCCCCCTGAGCGAG   1476
Db      1450 AAGGAGCTGTACCCCTTGGCCTCCTGTAGGTCTCTGTGGCAACGACCTCTCCAG   1509
Qy      1477 TAA   1479
Db      1510 TAA   1512

RESULT 15
US-09-952-060-25
; Sequence 25, Application US/09952060 .
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Eminli, Emilio A.
; APPLICANT: Youll, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 37474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA encoding pMRKAd5 HIV-1 gag, coding
US-09-952-060-25

Query Match          66.0%; Score 976.6; DB 4; Length 37474;
Best Local Similarity 80.2%; Pred. No. 1,4e-142;
Matches 1205; Conservative 0; Mismatches 274; Indels 24; Gaps 4;

Qy      1 ATGGGCGCGCGCGCAGCATCTCGCGCGCGGCAAGCTGGA CGGCTGGAGCGCATCGC 60
Db      1272 ATGGGTGTTAAGGCTTCTGTCTGTCTGTGTGTGAGCTTGACAATGTGAGAGATCAGG 1331
Qy      61 CTGGCGCCCGCGCGCAAGATGTGTACTATGTGAAGCACTGTGTGTGGCCAGCCGCGAG 120
Db      1332 CTGAGGCGCTGTGTGCAAGAAAGATAAGTAACAGCTTAAGCACTTGTGTGGGCTCTCAGGGAG 1391
Qy      121 CTGAGAAAGTTTCGCCCTTAAACCCCGGCGCTGTGTGAAGACCAACGAGGCTGTGAACGATC 180
Db      1392 CTGAGAGAGTTTGTCTGTAAACCTTGCGCTGTGAAGACCTTCTGAGGGGTGTGAGCAATC 1451
Qy      181 ATCGCCACAGCTGCACCCCGCGCTGTGAAGCGGCAAGCGAGAGCTGAAGAGCTGTGTAAC 240
Db      1452 CTGGGCGACGCTTCAGCCCTCTCTGCAAAAGAGCTCTGAAGAGCTGAGGTTCTCTTACAC 1511
Qy      241 ACCGTGGCAACCTGTACTGTGCTGCACGAAAGATCAGAGTTCGCGCACCAACGAGAGGCC 300

```

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 21:10:07 ; Search time 917.198 Seconds  
(without alignments)  
10560.254 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 1479  
Sequence: 1 atgggcgcgcgcgcgcagcat.....acgaccctcgtgacagctaa 1479

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10I\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1479	100.0	1479	10	US-09-967-464-63
2	1479	100.0	1479	10	US-09-899-575-3
3	1477.4	99.9	1479	15	US-10-190-435-18
4	1477.4	99.9	1479	15	US-10-190-435-16
5	1477.4	99.9	1479	15	US-10-190-435-19
6	1477.4	99.9	1479	15	US-10-190-305A-14
7	1477.4	99.9	1479	15	US-10-190-435-35

8	1477.4	99.9	4606	15	US-10-190-435-34	Sequence 34, Appl
9	1477.4	99.9	4615	15	US-10-190-435-36	Sequence 36, Appl
10	1477.4	99.9	4702	15	US-10-190-435-38	Sequence 38, Appl
11	1477.4	99.9	4716	15	US-10-190-435-17	Sequence 17, Appl
12	1477.4	99.9	4716	16	US-10-190-305A-13	Sequence 13, Appl
13	1475.8	99.8	2742	15	US-10-190-435-20	Sequence 20, Appl
14	1475.8	99.8	2742	16	US-10-190-305A-15	Sequence 15, Appl
15	1475.8	99.8	3330	15	US-10-190-435-9	Sequence 9, Appl
16	1475.8	99.8	3330	15	US-10-190-435-10	Sequence 10, Appl
17	1475.8	99.8	3330	15	US-10-190-435-11	Sequence 11, Appl
18	1475.8	99.8	5145	15	US-10-190-435-12	Sequence 12, Appl
19	1475.8	99.8	5145	16	US-10-190-305A-12	Sequence 12, Appl
20	1474.8	99.7	4713	15	US-10-190-435-17	Sequence 17, Appl
21	1474.8	99.7	4713	16	US-10-190-305A-83	Sequence 83, Appl
22	1473.8	99.6	5184	15	US-10-190-435-58	Sequence 58, Appl
23	1473.8	99.6	5184	16	US-10-190-305A-82	Sequence 82, Appl
24	1473.4	99.6	1479	10	US-09-967-464-67	Sequence 67, Appl
25	1472.4	99.6	2742	15	US-10-190-435-57	Sequence 57, Appl
26	1472.4	99.6	2742	16	US-10-190-305A-81	Sequence 81, Appl
27	1463	98.9	1479	10	US-09-899-575-20	Sequence 20, Appl
28	1315.2	88.9	3531	15	US-10-190-435-13	Sequence 13, Appl
29	1315.2	88.9	3537	15	US-10-190-435-14	Sequence 14, Appl
30	1315.2	88.9	3537	15	US-10-190-435-15	Sequence 15, Appl
31	1288.8	87.1	1491	10	US-09-899-575-51	Sequence 51, Appl
32	1288	87.1	1491	10	US-09-899-575-51	Sequence 51, Appl
33	1276.8	86.3	1509	10	US-09-899-575-4	Sequence 4, Appl
34	1275.2	86.2	1509	10	US-09-967-464-64	Sequence 64, Appl
35	1273.6	86.1	1509	10	US-09-967-464-68	Sequence 68, Appl
36	1260.8	85.2	1509	10	US-09-899-575-21	Sequence 21, Appl
37	1221.2	82.6	9166	18	US-10-359-120-168	Sequence 168, App
38	1206	81.5	9788	18	US-10-359-120-174	Sequence 174, App
39	1197.8	81.0	1515	17	US-10-387-336-4	Sequence 4, Appl
40	1197.8	81.0	4472	17	US-10-387-336-75	Sequence 75, Appl
41	1197.8	81.0	4608	17	US-10-387-336-76	Sequence 76, Appl
42	1197.8	81.0	4689	17	US-10-387-336-74	Sequence 74, Appl
43	1197.8	81.0	4766	17	US-10-387-336-73	Sequence 73, Appl
44	1196.2	80.9	2799	16	US-10-241-009-18	Sequence 18, Appl
45	1196.2	80.9	2799	16	US-10-190-434B-18	Sequence 18, Appl

#### ALIGNMENTS

RESULT 1  
US-09-967-464-63  
; Sequence 63, Application US/09967464  
; Publication No. US20030138453A1  
GENERAL INFORMATION:  
; APPLICANT: O'Hagan, Derek  
; APPLICANT: Otten, Gillie  
; APPLICANT: Donnelly, John J.  
; APPLICANT: Polo, John M.  
; APPLICANT: Barnett, Susan  
; APPLICANT: Singh, Manohan  
; APPLICANT: Ulmer, Jeffrey  
; APPLICANT: Dubensky, Jr., Thomas W.  
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS  
; FILE REFERENCE: P16269.004  
; CURRENT APPLICATION NUMBER: US/09/967,464  
; PRIOR FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 60/236,105  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/315,905  
; PRIOR FILING DATE: 2001-08-30  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 63  
; LENGTH: 1479  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-967-464-63  
Query Match 100.0%; Score 1479; DB 10; Length 1479;

Best Local Similarity 100.0%; Pred. No. 4.6e-308;  
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCCGGCGCAGCATCTGGCGCGCGCGCAAGCTGGAGCGCTGGGAGGGCATCCGC 60
Db 1 ATGGGCGCCGGCGCAGCATCTGGCGCGCGCGCAAGCTGGAGCGCTGGGAGGGCATCCGC 60
QY 61 CTGGCGCCGGCGCGCAAGAGTGTCTACATGATGAAGACCTGTGTGGCCAGCCGCGAG 120
Db 61 CTGGCGCCGGCGCGCAAGAGTGTCTACATGATGAAGACCTGTGTGGCCAGCCGCGAG 120
QY 121 CTGGAAGAGTTGGCTTGAACCCCGGCTGTGGAAGACGAGGAGGCTGCAAGCATC 180
Db 121 CTGGAAGAGTTGGCTTGAACCCCGGCTGTGGAAGACGAGGAGGCTGCAAGCATC 180
QY 181 ATCCGCGACGTGACCCCGGCTGCAAGACGCGGAGGAGAGCTGAAGAGCTGTTCAC 240
Db 181 ATCCGCGACGTGACCCCGGCTGCAAGACGCGGAGGAGAGCTGAAGAGCTGTTCAC 240
QY 241 ACCGTGGCACCCTGTACTGCTGCAAGAGATGAGTCCGCGACCAAGAGGCG 300
Db 241 ACCGTGGCACCCTGTACTGCTGCAAGAGATGAGTCCGCGACCAAGAGGCG 300
QY 301 CTGGAACAAGATGAGGAGGAGCAAGAACTGTCAGCAGAGATCCAGAGGCGC 360
Db 301 CTGGAACAAGATGAGGAGGAGCAAGAACTGTCAGCAGAGATCCAGAGGCGC 360
QY 361 GCGGCAAGAGGAGGAGTGAAGGAGCACTACCCATGTGTCAGAACTGTGAGGCGCAG 420
Db 361 GCGGCAAGAGGAGGAGTGAAGGAGCACTACCCATGTGTCAGAACTGTGAGGCGCAG 420
QY 421 GTGCAACAGAGCATGAGCCCGCGCACCTGTAACGCTGTGATGAAGTATCGAGAG 480
Db 421 GTGCAACAGAGCATGAGCCCGCGCACCTGTAACGCTGTGATGAAGTATCGAGAG 480
QY 481 GCTTTCAGCCCGGAGGTGATCCCATGTTCAACGCTGAGAGGAGGCGCCACCCCG 540
Db 481 GCTTTCAGCCCGGAGGTGATCCCATGTTCAACGCTGAGAGGAGGCGCCACCCCG 540
QY 541 GACCTGAACAGAGTGTGAACAACCGTGGGCGCGCACAGCGCGCATGCAATGCTGA 600
Db 541 GACCTGAACAGAGTGTGAACAACCGTGGGCGCGCACAGCGCGCATGCAATGCTGA 600
QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCGCTGACCGCGCG 660
Db 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCGTGGAGCCGCGTGCACCGCG 660
QY 661 ATGCGCCCGGCGCAAGTGGCGAGCGCGCGCGAGCATGTCGCGGACCAACAGCACC 720
Db 661 ATGCGCCCGGCGCAAGTGGCGAGCGCGCGCGAGCATGTCGCGGACCAACAGCACC 720
QY 721 CTGGAAGAGCATGCGCTGGATGAACAGCAACCCCGCATCCGCTGGGCGCATCTAC 780
Db 721 CTGGAAGAGCATGCGCTGGATGAACAGCAACCCCGCATCCGCTGGGCGCATCTAC 780
QY 781 AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGACAGCCCGTGAAGATC 840
Db 781 AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGACAGCCCGTGAAGATC 840
QY 841 CTGGAACATTAAGAGAGGCGCCCAAGAGGCGCTTCCGAGCTACGTGAGCCGCTTCA 900
Db 841 CTGGAACATTAAGAGAGGCGCCCAAGAGGCGCTTCCGAGCTACGTGAGCCGCTTCA 900
QY 901 ACCCTGGCGCGGAGAGAGCAACCAAGAGTGAAGAACTGATGACGAGCAACCCGCG 960
Db 901 ACCCTGGCGCGGAGAGAGCAACCAAGAGTGAAGAACTGATGACGAGCAACCCGCG 960
QY 961 GTGGAAGAGCGCAACCCCGAGCTGCAAGACATCTGCGCGCTCTGCGCCCGCGCGAG 1020
Db 961 GTGGAAGAGCGCAACCCCGAGCTGCAAGACATCTGCGCGCTCTGCGCCCGCGCGAG 1020
QY 1021 CTGGAAGAGATGATGACCGCTGCGAGGCGTGGGCGGCGCCCAACCAAGGCGCGGTG 1080
Db 1021 CTGGAAGAGATGATGATGACCGCTGCGAGGCGTGGGCGGCGCCCAACCAAGGCGCGGTG 1080
```

```
Db 1021 CTGGAAGAGATGATGATGACCGCTGCGAGGCGTGGGCGGCGCCCAACCAAGGCGCGGTG 1080
QY 1081 CTGGCGGAGGAGATGAGCGCAGGCCCAACACGACCGTATGATGAGAGAGCACTTCAAG 1140
Db 1081 CTGGCGGAGGAGATGAGCGCAGGCCCAACACGACCGTATGATGAGAGAGCACTTCAAG 1140
QY 1141 GCGCCCGGCGCATCTGTCAGATGCTTCAACTGCGGCAAGAGAGGCCCATGCGCGCAAC 1200
Db 1141 GCGCCCGGCGCATCTGTCAGATGCTTCAACTGCGGCAAGAGAGGCCCATGCGCGCAAC 1200
QY 1201 TGCGCGCGCCCGCGCAAGAGGCTGTGGAAGTGCAGAGAGGAGGCGCACATGTAAG 1260
Db 1201 TGCGCGCGCCCGCGCAAGAGGCTGTGGAAGTGCAGAGAGGAGGCGCACATGTAAG 1260
QY 1261 GACTGCAACGAGCGCGCAGGCCCACTTCTGGGCAAGATCTGGCCAGCCAGCAAGAGCGCG 1320
Db 1261 GACTGCAACGAGCGCGCAGGCCCACTTCTGGGCAAGATCTGGCCAGCCAGCAAGAGCGCG 1320
QY 1321 CCGGCAACTTCTGTCAGAGCGCGCCGAGGCCACCGCCCGCGCGAGAGCTTCCGC 1380
Db 1321 CCGGCAACTTCTGTCAGAGCGCGCCGAGGCCACCGCCCGCGCGAGAGCTTCCGC 1380
QY 1381 TTGGAAGAGACCAACCCCGCGCAAGAGCAGAGAGCAAGAGCCGAGAGCTTGAACAGC 1440
Db 1381 TTGGAAGAGACCAACCCCGCGCAAGAGCAGAGAGCAAGAGCCGAGAGCTTGAACAGC 1440
QY 1441 CTGAAGAGCTGTTCGCGCAAGCAACCCCTGAGCCAGTAA 1479
Db 1441 CTGAAGAGCTGTTCGCGCAAGCAACCCCTGAGCCAGTAA 1479
```

```
RESULT 2
US-09-899-575-3
; Sequence 3, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megele, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Esterlita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USBS THEROF
; FILE REFERENCE: P01631.102
; CURRENT FILING DATE: 2001-07-05
; PRIOR FILING DATE: 1999-12-30
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110965
US-09-899-575-3
```

```
Query Match 100.0%; Score 1479; DB 10; Length 1479;
Best Local Similarity 100.0%; Pred. No. 4.6e-308;
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGCCGGCGCAGCATCTGGCGCGCGCGCAAGCTGGAGCGCTGGGAGGGCATCCGC 60
Db 1 ATGGGCGCCGGCGCAGCATCTGGCGCGCGCGCAAGCTGGAGCGCTGGGAGGGCATCCGC 60
QY 61 CTGGCGCCGGCGCGCAAGAGTGTCTACATGATGAAGACCTGTGTGGCCAGCCGCGAG 120
Db 61 CTGGCGCCGGCGCGCAAGAGTGTCTACATGATGAAGACCTGTGTGGCCAGCCGCGAG 120
QY 121 CTGGAAGAGTTGGCTTGAACCCCGGCTGTGGAAGACGAGGAGGCTGCAAGCATC 180
Db 121 CTGGAAGAGTTGGCTTGAACCCCGGCTGTGGAAGACGAGGAGGCTGCAAGCATC 180
```







```

QY 661 ATGCCCCCGCCAGATGGCGGAGCCCGCGGCGAGCATGGCCGACACACAGACC 720
    |||
    |||
    |||
Db 667 ATGCCCCCGCCAGATGGCGGAGCCCGCGGCGAGCATGGCCGACACACAGACC 726
    |||
    |||
    |||
QY 721 CTGCAAGAGCAGATCGCTGGATGACAGCAACCCCGCATCCCGTGGCGACATCTAC 780
    |||
    |||
    |||
Db 727 CTGCAAGAGCAGATCGCTGGATGACAGCAACCCCGCATCCCGTGGCGAGCATCTAC 786
    |||
    |||
    |||
QY 781 AAGCGGTGATCATCTGGGCTTGAACAAAGATGTGCGAGATGTACAGCCCGTAGCATC 840
    |||
    |||
    |||
Db 787 AAGCGGTGATCATCTGGGCTTGAACAAAGATGTGCGAGATGTACAGCCCGTAGCATC 846
    |||
    |||
    |||
QY 841 CTGGAATCAAGCAGGGGCCCAAGAGCCCTTCCGCACTAGCTGAGCCGCTTTCAAG 900
    |||
    |||
    |||
Db 847 CTGGAATCAAGCAGGGGCCCAAGAGCCCTTCCGCACTAGCTGAGCCGCTTTCAAG 906
    |||
    |||
    |||
QY 901 ACCCTGCGCGCGAGAGAGCAACCCAGAGGTGAAGAACTGGATGACGAGACCCCTGCTG 960
    |||
    |||
    |||
Db 907 ACCCTGCGCGCGAGAGAGCAACCCAGAGGTGAAGAACTGGATGACGAGACCCCTGCTG 966
    |||
    |||
    |||
QY 961 GTGCAGAACGCCAACCCCGCATCTGCAGAACCATCTGCGGCTCTCGGCCCGCGCGCAC 1020
    |||
    |||
    |||
Db 967 GTGCAGAACGCCAACCCCGCATCTGCAGAACCATCTGCGGCTCTCGGCCCGCGCGCAC 1026
    |||
    |||
    |||
QY 1021 CTGGAAGAGATGATGACCCGCTGCGCAGGGCGTGGGCGGCCCAAGCCAGCCGCGCTG 1080
    |||
    |||
    |||
Db 1027 CTGGAAGAGATGATGACCCGCTGCGCAGGGCGTGGGCGGCCCAAGCCAGCCGCGCTG 1086
    |||
    |||
    |||
QY 1081 CTGGCGAGAGCATGAGCAGAGGCAACACAGCGTGAATGATGAGAAGCAACTTCAAG 1140
    |||
    |||
    |||
Db 1087 CTGGCGAGAGCATGAGCAGAGGCAACACAGCGTGAATGATGAGAAGCAACTTCAAG 1146
    |||
    |||
    |||
QY 1141 GGGCCCCGCGCATCGTCAAGTGTTCATCTGCGGCAAGAGGCGCCATCCCGCGCAAC 1200
    |||
    |||
    |||
Db 1147 GGGCCCCGCGCATCGTCAAGTGTTCATCTGCGGCAAGAGGCGCCATCCCGCGCAAC 1206
    |||
    |||
    |||
QY 1201 TGGCGGCGCCCCCGCAAGAGGCTGTGGAAGTGTGCGCAAGAGGCGCCATCAATGAAG 1260
    |||
    |||
    |||
Db 1207 TGGCGGCGCCCCCGCAAGAGGCTGTGGAAGTGTGCGCAAGAGGCGCCATCAATGAAG 1266
    |||
    |||
    |||
QY 1261 GACTGACCGAGCGCAGAGGCAACTTCCGGGCAAGATGTGGCCCAAGAGGCGCGC 1320
    |||
    |||
    |||
Db 1267 GACTGACCGAGCGCAGAGGCAACTTCCGGGCAAGATGTGGCCCAAGAGGCGCGC 1326
    |||
    |||
    |||
QY 1321 CCGGCAACTTCTGTGAGAGCGCCCGGAGCCCAAGCCCGCCCGCGAGAGCTTCCGC 1380
    |||
    |||
    |||
Db 1327 CCGGCAACTTCTGTGAGAGCGCCCGGAGCCCAAGCCCGCCCGCGAGAGCTTCCGC 1386
    |||
    |||
    |||
QY 1381 TTGCAAGAGACCAACCCCGCGCAAGAGCAGAGAGCAGAGCCGCGAGACCTTGAACAGC 1440
    |||
    |||
    |||
Db 1387 TTGCAAGAGACCAACCCCGCGCAAGAGCAGAGAGCAGAGCCGCGAGACCTTGAACAGC 1446
    |||
    |||
    |||
QY 1441 CTGAAGAGCGTGTGGGCAAGACCCCGTAGAGCAAGTAA 1479
    |||
    |||
    |||
Db 1447 CTGAAGAGCGTGTGGGCAAGACCCCGTAGAGCAAGTAA 1485
    |||
    |||
    |||

```

```

RESULT 5
US-10-190-435-19
; Sequence 19, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrellita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P/18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30

```

```

; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmutatRevNef_C
US-10-190-435-19

```

```

Query Match      99.9%; Score 1477.4; DB 15; Length 4419;
Best Local Similarity 99.9%; Pred. No. 9.6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGGCGCCCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTTGGAGGCGATCCGC 60
    |||
    |||
    |||
Db 7 ATGGCGCCCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTTGGAGGCGATCCGC 66
    |||
    |||
    |||
QY 61 CTGCGCCCGCGCGCAAGAGTGTCTAATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
    |||
    |||
    |||
Db 67 CTGCGCCCGCGCGCAAGAGTGTCTAATGATGAAGCACTGTGTGGGCCAGCCGCGAG 126
    |||
    |||
    |||
QY 121 CTGGAAGATTGACCTGTAACCCCGGCTGTGAGAGCAGAGAGGCGTGCAGAGATC 180
    |||
    |||
    |||
Db 127 CTGGAAGATTGACCTGTAACCCCGGCTGTGAGAGCAGAGAGGCGTGCAGAGATC 186
    |||
    |||
    |||
QY 181 ATCCGCAAGCTGCAACCCCGCTGCGAGCCGCGAGAGAGAGCTGMAAGCCTGTTCAAC 240
    |||
    |||
    |||
Db 187 ATCCGCAAGCTGCAACCCCGCTGCGAGCCGCGAGAGAGAGCTGMAAGCCTGTTCAAC 246
    |||
    |||
    |||
QY 241 ACCGTGGCACCCCTGTACTGTGTGCAAGAAAGATCGAGTCCGCGACACCAAGAGGCC 300
    |||
    |||
    |||
Db 247 ACCGTGGCACCCCTGTACTGTGTGCAAGAAAGATCGAGTCCGCGACACCAAGAGGCC 306
    |||
    |||
    |||
QY 301 CTGGAAGATGAGGAAGAGAGAGCAAGATGCCAGAGAGATCAGAGGCGGAGGCC 360
    |||
    |||
    |||
Db 307 CTGGAAGATGAGGAAGAGAGAGCAAGATGCCAGAGAGATCAGAGGCGGAGGCC 366
    |||
    |||
    |||
QY 361 GCCGCAAGAGGCAAGGTGAGCAGAACTAACCTATCTGTGCAAACTTGCAGAGGCGCAGATG 420
    |||
    |||
    |||
Db 367 GCCGCAAGAGGCAAGGTGAGCAGAACTAACCTATCTGTGCAAACTTGCAGAGGCGCAGATG 426
    |||
    |||
    |||
QY 421 GTGCACCAAGGCTATGAGCCCGCGCAACCTTGAACCTTGGTGAAGGTATGAGAGAGAG 480
    |||
    |||
    |||
Db 427 GTGCACCAAGGCTATGAGCCCGCGCAACCTTGAACCTTGGTGAAGGTATGAGAGAGAG 486
    |||
    |||
    |||
QY 481 GCTTCAAGCCCGAGAGTATCCCATGTTCAACCGGCGTGAAGAGGCGGCCAACCCCGAG 540
    |||
    |||
    |||
Db 487 GCTTCAAGCCCGAGAGTATCCCATGTTCAACCGGCGTGAAGAGGCGGCCAACCCCGAG 546
    |||
    |||
    |||
QY 541 GACCTGAACAGAGTGTGAACACCGTGGCGCGCAACAGGCGCGCATGACAGATGCTGAAG 600
    |||
    |||
    |||
Db 547 GACCTGAACAGAGTGTGAACACCGTGGCGCGCAACAGGCGCGCATGACAGATGCTGAAG 606
    |||
    |||
    |||
QY 601 GACACATCAACAGAGAGCGCGCGAGTGGAGCGCGGTGACACCCCGTGCACGCGCGGCC 660
    |||
    |||
    |||
Db 607 GACACATCAACAGAGAGCGCGCGAGTGGAGCGCGGTGACACCCCGTGCACGCGCGGCC 666
    |||
    |||
    |||
QY 661 ATGCGCCCGCGCGAGATCGGAGAGCGCCGCGCGAGCATGCGCGGCAACCAAGAGACC 720
    |||
    |||
    |||
Db 667 ATGCGCCCGCGCGAGATCGGAGAGCGCCGCGCGAGCATGCGCGGCAACCAAGAGACC 726
    |||
    |||
    |||
QY 721 CTGCAAGAGCAGATCGCTGGATGACAGCAACCCCGCATCCCGTGGCGAGCATCTAC 780
    |||
    |||
    |||
Db 727 CTGCAAGAGCAGATCGCTGGATGACAGCAACCCCGCATCCCGTGGCGAGCATCTAC 786
    |||
    |||
    |||
QY 781 AAGCGGTGATCATCTGGGCTTGAACAAAGTGTGCGAGATGTACAGCCCGTGAAGCATC 840
    |||
    |||
    |||
Db 787 AAGCGGTGATCATCTGGGCTTGAACAAAGTGTGCGAGATGTACAGCCCGTGAAGCATC 846
    |||
    |||
    |||
QY 841 CTGGAATCAAGCAGGGGCCCAAGAGGCGCTTCCGCGCACTAGCTGAGCCGCTTTCAAG 900
    |||
    |||
    |||
Db 847 CTGGAATCAAGCAGGGGCCCAAGAGGCGCTTCCGCGCACTAGCTGAGCCGCTTTCAAG 906
    |||
    |||
    |||

```

QY 901 ACCCTGCGCGCGGAGAGAGCACCAGAGGTGAAGAACTGGATGACCGGACACCTTGCTG 960  
DB 907 ACCCTGCGCGCGGAGAGAGCACCAGAGGTGAAGAACTGGATGACCGGACACCTTGCTG 966  
QY 961 GTGCAAGAACGCAACCCCGCATCTGCAAGACATCTTGCGGCTCTGCGGCCCGCGCGCAGC 1020  
DB 967 GTGCAAGAACGCAACCCCGCATCTGCAAGACATCTTGCGGCTCTGCGGCCCGCGCGCAGC 1026  
QY 1021 CTGGAGAGATGATGACCCCTGCTGCGAGGGCGGTGGGGGCCCAAGCCCAAGGCCCGCGCTG 1080  
DB 1027 CTGGAGAGATGATGACCCCTGCTGCGAGGGCGGTGGGGGCCCAAGGCCCGCGCTG 1086  
QY 1081 CTGGCGAGAGCATGAGCCAGGCAACACACGCGTGATGATGCAGAGAGCACTTCAAG 1140  
DB 1087 CTGGCGAGAGCATGAGCCAGGCAACACACGCGTGATGATGCAGAGAGCACTTCAAG 1146  
QY 1141 GGCCCCCGGCGCATGCTCAAGTGTCTTCAACTGCGGCAAGAGAGGCCCATGCGCCGCAAC 1200  
DB 1147 GGCCCCCGGCGCATGCTCAAGTGTCTTCAACTGCGGCAAGAGAGGCCCATGCGCCGCAAC 1206  
QY 1201 TGGCGGCGCCCCCGCAAGAGGGGTCTGGAAGTGTGGGCAAGAGAGGCCCAACGATGAG 1260  
DB 1207 TGGCGGCGCCCCCGCAAGAGGGGTCTGGAAGTGTGGGCAAGAGAGGCCCAACGATGAG 1266  
QY 1261 GACTGACCGAGGCGCCAGGCAACTTCTTGCGCAAGATCTGGGCCAGCCACCAAGGGCGCG 1320  
DB 1267 GACTGACCGAGGCGCCAGGCAACTTCTTGCGCAAGATCTGGGCCAGCCACCAAGGGCGCG 1326  
QY 1321 CCGCGCACTTCTTGAGAGCGCGCCCGAGGCCCAAGCCGCCCGCGGAGAGCTTTCGCG 1380  
DB 1327 CCGCGCACTTCTTGAGAGCGCGCCCGAGGCCCAAGCCGCCCGCGGAGAGCTTTCGCG 1386  
QY 1381 TTTCGAGAGAACACCCCCCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
DB 1387 TTTCGAGAGAACACCCCCCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446  
QY 1441 CTGAAGAGCGCTGTTCGAGCAAGCAACCCCTGAGCCAGTAA 1479  
DB 1447 CTGAAGAGCGCTGTTCGAGCAAGCAACCCCTGAGCCAGTAA 1485

RESULT 6  
US-10-190-305A-14  
; Sequence 14, Application US/10190305A  
; Publication No. US20030198621A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARNETT, Susan  
; APPLICANT: LIAN, Ying  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR  
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 2302-18702 / 18702.002  
; CURRENT APPLICATION NUMBER: US/10/190,305A  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 4419  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: GagRtmtrtatRevNef\_C  
US-10-190-305A-14

Query Match 99.9%; Score 1477.4; DB 16; Length 4419;  
Best Local Similarity 99.9%; Pred. No. 9.6e-308;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGGCGAGCTGAGCGCTGGAGAGCGATCCGC 60  
DB 7 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGGCGAGCTGAGCGCTGGAGAGCGATCCGC 66

QY 61 CTGCGCGCGCGGCAAGAGTGCTACATGATGAAGCACTGTGTGTGGCGCAGCGCAG 120  
DB 67 CTGCGCGCGCGGCAAGAGTGCTACATGATGAAGCACTGTGTGTGGCGCAGCGCAG 126  
QY 121 CTGAGAGAGTTTCGCTTGAACCCCGCTGTGAGAGACAGCGAGGCTGCAAGCATC 180  
DB 127 CTGAGAGAGTTTCGCTTGAACCCCGCTGTGAGAGACAGCGAGGCTGCAAGCATC 186  
QY 181 ATCCGCGAGCTGACACCCCGCTTCAAGCCGCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 187 ATCCGCGAGCTGACACCCCGCTTCAAGCCGCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 246  
QY 241 ACCGTGGCACCTGTATCTGCTGCAACGAGAGATGAGTGTCCGCAACCAAGAGAGCC 300  
DB 247 ACCGTGGCACCTGTATCTGCTGCAACGAGAGATGAGTGTCCGCAACCAAGAGAGCC 306  
QY 301 CTGAGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 307 CTGAGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366  
QY 361 GCCGCAAGGGCAAGGTGAGAGCAACCTACCTGCGAGAACTGCGAGAGGCGCAGATG 420  
DB 367 GCCGCAAGGGCAAGGTGAGAGCAACCTACCTGCGAGAACTGCGAGAGGCGCAGATG 426  
QY 421 GTGCAACAGGCGCATCAGCGCCCGCAACCTGAAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
DB 427 GTGCAACAGGCGCATCAGCGCCCGCAACCTGAAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 486  
QY 481 GCTTTACGCGCGAGAGTATCTCCATGTTTCAACCGCGCTGAGCGAGAGGCGCAGCCCGCAG 540  
DB 487 GCTTTACGCGCGAGAGTATCTCCATGTTTCAACCGCGCTGAGCGAGAGGCGCAGCCCGCAG 546  
QY 541 GACCTGAACAGAGTGTGAACACGCTGGGGGGGCAACAGCGCGGCAATGACATGCTGAAG 600  
DB 547 GACCTGAACAGAGTGTGAACACGCTGGGGGGGCAACAGCGCGGCAATGACATGCTGAAG 606  
QY 601 GACACCATCAACGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
DB 607 GACACCATCAACGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666  
QY 661 ATGCGCGCGCGAGAGTCCGAGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
DB 667 ATGCGCGCGCGAGAGTCCGAGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726  
QY 721 CTGCAAGAGAGATGCTGCTGATGACAGCAACCCCGCATCCCGTGGGGAGATCTAC 780  
DB 727 CTGCAAGAGAGATGCTGCTGATGACAGCAACCCCGCATCCCGTGGGGAGATCTAC 786  
QY 781 AAGCGGTGATCATCTGGGCTTGAACAGATGCTGCGATGTACAGCCCGGTGAGCATC 840  
DB 787 AAGCGGTGATCATCTGGGCTTGAACAGATGCTGCGATGTACAGCCCGGTGAGCATC 846  
QY 841 CTGAGCATCAAGAGAGCGCGCGCAAGAGAGCGCTTTCGCGCATGAGAGAGAGAGAGAGAGAGAGAGAG 900  
DB 847 CTGAGCATCAAGAGAGCGCGCGCAAGAGAGCGCTTTCGCGCATGAGAGAGAGAGAGAGAGAGAGAGAG 906  
QY 901 ACCCTGCGCGCGGAGAGAGCAACCAAGAGAGTGAAGAACTGGATGACCGGACACCTTGCTG 960  
DB 907 ACCCTGCGCGCGGAGAGAGCAACCAAGAGAGTGAAGAACTGGATGACCGGACACCTTGCTG 966  
QY 961 GTGCAAGAGCGCAACCCCGCATCTGCAAGACATCTTGCGGCTCTGCGGCCCGCGCGCAGC 1020  
DB 967 GTGCAAGAGCGCAACCCCGCATCTGCAAGACATCTTGCGGCTCTGCGGCCCGCGCGCAGC 1026  
QY 1021 CTGGAGAGATGATGACCGCTGCTGCGAGGGCGGTGGGGGCCCAAGGCCCGCGCTG 1080  
DB 1027 CTGGAGAGATGATGACCGCTGCTGCGAGGGCGGTGGGGGCCCAAGGCCCGCGCTG 1086  
QY 1081 CTGGCGAGAGCATGAGCCAGGCAACACACGCGTGATGATGCAGAGAGCACTTCAAG 1140  
DB 1087 CTGGCGAGAGCATGAGCCAGGCAACACACGCGTGATGATGCAGAGAGCACTTCAAG 1146

QY 1141 GGCCCCGGCGCATCTCAAGTGTCTTCAACTGCGCGCAAGAGGGCCACATCGCCCGCAAC 1200  
|  
|  
|  
Db 1147 GGCCCCGGCGCATCTCAAGTGTCTTCAACTGCGCGCAAGAGGGCCACATCGCCCGCAAC 1206  
|  
|  
|  
QY 1201 TGGCGGCGCGCGCGCAAGAGGGCTGTGGAAGTGTGCGCAAGAGGGCCACAGATGAAG 1260  
|  
|  
|  
Db 1207 TGGCGGCGCGCGCGCAAGAGGGCTGTGGAAGTGTGCGCAAGAGGGCCACAGATGAAG 1266  
|  
|  
|  
QY 1261 GACTGACCGAGCGCGCAAGGGCCCACTTCTGTGGCGCAAGATGTGGCCCGCAAGAGGGCCGC 1320  
|  
|  
|  
Db 1267 GACTGACCGAGCGCGCAAGGGCCCACTTCTGTGGCGCAAGATGTGGCCCGCAAGAGGGCCGC 1326  
|  
|  
|  
QY 1321 CCGCGCACTTCTGTGAGAGCGCGCGAGCCCAAGCGCCCGCGCGAGAGCTTCCGC 1380  
|  
|  
|  
Db 1327 CCGCGCACTTCTGTGAGAGCGCGCGAGCCCAAGCGCCCGCGCGAGAGCTTCCGC 1386  
|  
|  
|  
QY 1381 TTGAGAGAGCAACCCCGCGCGCAAGAGCAAGAGAGCAAGAGCGCGAGACCTTGAACAGC 1440  
|  
|  
|  
Db 1387 TTGAGAGAGCAACCCCGCGCGCAAGAGCAAGAGAGCAAGAGCGCGAGACCTTGAACAGC 1446  
|  
|  
|  
QY 1441 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479  
|  
|  
|  
Db 1447 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1485  
|  
|  
|

## RESULT 7

US-10-190-435-35  
; Sequence 35, Application US/10190435  
; Publication No. US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Estrelita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: P/8133.003 / 2302-18133  
; CURRENT APPLICATION NUMBER: US/10/190,435  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: Patent In Ver. 2.0  
; LENGTH: 4483  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: gp160mod.TVI.dv1V2-gagmod.BM965  
US-10-190-435-35

Query Match 99.9%; Score 1477.4; DB 15; Length 4483;  
Best Local Similarity 99.9%; Pred. No. 9.6e-308;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGCGCAAGATCTCTGCGCGCGCGCAAGCTGAGCTCTGGAGGCGATCCGC 60  
|  
|  
|  
Db 3005 ATGGGCGCGCGCGCGCAAGATCTCTGCGCGCGCGCAAGCTGAGCTCTGGAGGCGATCCGC 3064  
|  
|  
|  
QY 61 CTGGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGACCTGTGTGTGGCGAGCGCGAG 120  
|  
|  
|  
Db 3065 CTGGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGACCTGTGTGTGGCGAGCGCGAG 3124  
|  
|  
|  
QY 121 CTGGAAGATTGCTGTGAACCCCGGCTGTGTGAGACAGGAGGAGCTGCAAGAGATC 180  
|  
|  
|  
Db 3125 CTGGAAGATTGCTGTGAACCCCGGCTGTGTGAGACAGGAGGAGCTGCAAGAGATC 3184  
|  
|  
|  
QY 181 ATCCGCAAGTCAACCCCGGCTGTGAGACAGGAGGAGAGTGAAGAGCTTGTTCAC 240  
|  
|  
|  
Db 3185 ATCCGCAAGTCAACCCCGGCTGTGAGACAGGAGGAGAGTGAAGAGCTTGTTCAC 3244  
|  
|  
|  
QY 241 ACCGTGCGCAACCTGTACTGTGCTGACGAGAAGATGAGTCCGCGACCAAGAGAGCC 300  
|  
|  
|

Db 3245 ACCGTGCGCAACCTGTACTGTGCTGACGAGAAGATCGAGGTCCGCGACCAAGAGAGCC 3304  
|  
|  
|  
QY 301 CTGGAAGATCGAGAGAGAGCAACAGTGTCCAGCAAGAGATCAGAGGCGGAGCC 360  
|  
|  
|  
Db 3305 CTGGAAGAGATCGAGAGAGAGCAACAGTGTCCAGCAAGAGATCAGAGGCGGAGCC 3364  
|  
|  
|  
QY 361 GCCGCAAGGGGCAAGGTGAGCGAGAACTAACCCATCTGTGCAAGATCTTCAGGGCCAGATG 420  
|  
|  
|  
Db 3365 GCCGCAAGGGGCAAGGTGAGCGAGAACTAACCCATCTGTGCAAGATCTTCAGGGCCAGATG 3424  
|  
|  
|  
QY 421 GTGCAACAGGCGCATAGCGCCCGCAACCTTGAACGCTGTGTAAGTGTGAGAGAGAAG 480  
|  
|  
|  
Db 3425 GTGCAACAGGCGCATAGCGCCCGCAACCTTGAACGCTGTGTAAGTGTGAGAGAGAAG 3484  
|  
|  
|  
QY 481 GCTTTCAGCGCGAGGTGATCCCATGTTCAACCGGCTGTGAGAGAGGCGCAACCCCGAG 540  
|  
|  
|  
Db 3485 GCTTTCAGCGCGAGGTGATCCCATGTTCAACCGGCTGTGAGAGAGGCGCAACCCCGAG 3544  
|  
|  
|  
QY 541 GACCTGAACAGAGTGTGAACAACGTGTGGCGGCGCAACGAGCGCGCATGACAGATGCTGAAG 600  
|  
|  
|  
Db 3545 GACCTGAACAGAGTGTGAACAACGTGTGGCGGCGCAACGAGCGCGCATGACAGATGCTGAAG 3604  
|  
|  
|  
QY 601 GACACCATCAACGAGAGAGCGCGCGAGTGTGAGACCGCGTGCACCCCGTGCAGCGCGCC 660  
|  
|  
|  
Db 3605 GACACCATCAACGAGAGAGCGCGCGAGTGTGAGACCGCGTGCACCCCGTGCAGCGCGCC 3664  
|  
|  
|  
QY 661 ATGCGCCCGCGCGAGATCGGAGCGCGAGCGCGCGAGCGCATCGCGGCAACCAAGCAC 720  
|  
|  
|  
Db 3665 ATGCGCCCGCGCGAGATCGGAGCGCGAGCGCGCGAGCGCATCGCGGCAACCAAGCAC 3724  
|  
|  
|  
QY 721 CTGCAAGAGAGCATGTGCGCTGTGATGACGAGCAACCCCGCATCCCGTGTGGGAGCATTTAC 780  
|  
|  
|  
Db 3725 CTGCAAGAGAGCATGTGCGCTGTGATGACGAGCAACCCCGCATCCCGTGTGGGAGCATTTAC 3784  
|  
|  
|  
QY 781 AAGCGGTGATCATCTGTGGCTGTGAACAAGATGTGCGAGATGTACAGCCCGGTGAGCATC 840  
|  
|  
|  
Db 3785 AAGCGGTGATCATCTGTGGCTGTGAACAAGATGTGCGAGATGTACAGCCCGGTGAGCATC 3844  
|  
|  
|  
QY 841 CTGACATCAACGAGAGCGCCCAAGAGCGCTTTCGCGCATACGTGACCGGCTTTCAG 900  
|  
|  
|  
Db 3845 CTGACATCAACGAGAGCGCCCAAGAGCGCTTTCGCGCATACGTGACCGGCTTTCAG 3904  
|  
|  
|  
QY 901 ACCCTGCGCGCGAGAGAGCAACCAAGAGTGAAGATGAGACCGACACCTGTG 960  
|  
|  
|  
Db 3905 ACCCTGCGCGCGAGAGAGCAACCAAGAGTGAAGATGAGACCGACACCTGTG 3964  
|  
|  
|  
QY 961 GTGCAAGAGCGCAACCCCGCATGTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGAGC 1020  
|  
|  
|  
Db 3965 GTGCAAGAGCGCAACCCCGCATGTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGAGC 4024  
|  
|  
|  
QY 1021 CTGGAAGATGATGACCGCTGTGCGAGGCGGTGGGCGGCGCCAGCCCAAGGCGCGGTG 1080  
|  
|  
|  
Db 4025 CTGGAAGATGATGACCGCTGTGCGAGGCGGTGGGCGGCGCCAGCCCAAGGCGCGGTG 4084  
|  
|  
|  
QY 1081 CTGGCGGAGGAGTGTGAGCGCAAGCAACACAGCGTGTGATGTGCAAGAGCACTTCAAG 1140  
|  
|  
|  
Db 4085 CTGGCGGAGGAGTGTGAGCGCAAGCAACACAGCGTGTGATGTGCAAGAGCACTTCAAG 4144  
|  
|  
|  
QY 1141 GAGCGCGCGCGCATGTCAAGTGTCTCAATGCGCGCAAGAGAGGCAATGCGCGCAAC 1200  
|  
|  
|  
Db 4145 GAGCGCGCGCGCATGTCAAGTGTCTCAATGCGCGCAAGAGAGGCAATGCGCGCAAC 4204  
|  
|  
|  
QY 1201 TGGCGGCGCGCGCGCAAGAGGCTGTGAAGTGTGCGCAAGAGAGGCGCACAGATGAAG 1260  
|  
|  
|  
Db 4205 TGGCGGCGCGCGCGCAAGAGGCTGTGAAGTGTGCGCAAGAGAGGCGCACAGATGAAG 4264  
|  
|  
|  
QY 1261 GACTGACCGAGCGCGCAAGCTTCTGTGGCAAGATGTGGCCCAAGCAACAAGGCGCG 1320  
|  
|  
|  
Db 4265 GACTGACCGAGCGCGCAAGCTTCTGTGGCAAGATGTGGCCCAAGCAACAAGGCGCG 4324  
|  
|  
|  
QY 1321 CCGGCAACTTCTGTGAGAGCGCGCGAGCGCCACCGCGCGCGCGAGAGCTTCCGC 1380  
|  
|  
|  
Db 4325 CCGGCAACTTCTGTGAGAGCGCGCGAGCGCCACCGCGCGCGCGAGAGCTTCCGC 4384  
|  
|  
|



Query Match	99.9%	Score 1477.4;	DB 15;	Length 4615;
Best Local Similarity	99.9%;	Pred. No. 9.6e-308;		
Matches 1478; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	ATGGGCGCCCGCGCCAGATCTCTGGCGCGCGCAAGCTGAGCGCTTGGAGGCGATCCCG	60
Db	3137	ATGGGCGCGCCCGCGCGAGATCTTGGCGCGCGCAAGCTGAGCGCTTGGAGGCGATCCCG	3196
QY	61	CTGGCGCCCGCGCGCAAGAGTGTCTCATGTATGAAGCATCTGATGTGGGCCAGCGCGGAG	120
Db	3197	CTGGCGCCCGCGCGCGAGAGTGTCTCATGTATGAAGCATCTGATGTGGGCCAGCGCGGAG	3256
QY	121	CTGGAGAAATTGCGCTCTGAACCCCGGCTGTGTGAGACCAAGCGAGGTGCAAGCATTC	180
Db	3257	CTGGAGAAATTGCGCTCTGAACCCCGGCTGTGTGAGACCAAGCGAGGTGCAAGCATTC	3316
QY	181	ATCGCGCAGCTGCAACCCCGGCTGTGAGACCGGCGCGAGAGCTGAAAGCTTGTCAAC	240
Db	3317	ATCGCGCAGCTGCAACCCCGGCTGTGAGACCGGCGCGAGAGCTGAAAGCTTGTCAAC	3376
QY	241	ACCGTGGCCACTCTGTACTGTGCTGCACAGAAAGATCGAGTCTCGCGCACCAAGAGAGCC	300
Db	3377	ACCGTGGCCACTCTGTACTGTGCTGCACAGAAAGATCGAGTCTCGCGCACCAAGAGAGCC	3436
QY	301	CTGGACAAAGATTCGAGGAGAGAGACAGAACAAATGCCAGCAAGATTCAGCAGGCGGAGGCC	360
Db	3437	CTGGACAAAGATTCGAGGAGAGAGACAGAACAAATGCCAGCAAGATTCAGCAGGCGGAGGCC	3496
QY	361	GCCGACAAAGGCGCAGTGTGCGCAGAACTAACCCCTGTGTGAGAACTGTGAGAGCGACATG	420
Db	3497	GCCGACAAAGGCGCAGTGTGCGCAGAACTAACCCCTGTGTGAGAACTGTGAGAGCGCAGATG	3556
QY	421	GTCGACCAAGGCTCATGAGCCCGCGCACCTTGAACTGTGTGGTGAAGGTGATGAGAGAGAG	480
Db	3557	GTCGACCAAGGCTCATGAGCCCGCGCACCTTGAACTGTGTGGTGAAGGTGATGAGAGAGAG	3616
QY	481	GCCTTCAAGCCCGCAGGTGATCCCATGATTCAACCGCGCTGAGCGAGGGGCGCAACCCCGAG	540
Db	3517	GCCTTCAAGCCCGCAGGTGATCCCATGATTCAACCGCGCTGAGCGAGGGGCGCAACCCCGAG	3676
QY	541	GACCTGAACACGATGTTGAACACCGTGTGGCGGACCAAGGCGCGCATGACATGCTGAAG	600
Db	3677	GACCTGAACACGATGTTGAACACCGTGTGGCGGACCAAGGCGCGCATGACATGCTGAAG	3736
QY	601	GACACCATTCACGAGGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCAGCGCGGCCCC	660
Db	3737	GACACCATTCACGAGGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCAGCGCGGCCCC	3796
QY	661	ATGCGCCCGCGCGCAGATGCGGAGCCCGCGGACGACATGCGCGGACCAACGACGACCC	720
Db	3797	ATGCGCCCGCGCGCAGATGCGGAGCCCGCGGACGACATGCGCGGACCAACGACGACCC	3856

Qy	721	CTGCAGAGCAGATCGCTTGATGACCAAGCAACCCCCCATCCCGTGGGCAATCTAC	780
Db	3857	CTGCAGAGCAGATCGCTTGATGACCAAGCAACCCCCCATCCCGTGGGCAATCTAC	3916
Qy	781	AAGCGGTGATCATCTCTGGGCTTGAACAAATGTTGGGATTTACAGCCCCGTGAGCATC	840
Db	3917	AAGCGGTGATCATCTCTGGGCTTGAACAAATGTTGGGATTTACAGCCCCGTGAGCATC	3976
Qy	841	CTGCACATCAAGCAGGGCCCCCAAGAGAGCCCTTCGCGCATACGTTGACGCTTCTTCAAG	900
Db	3977	CTGCACATCAAGCAGGGCCCCCAAGAGAGCCCTTCGCGCATACGTTGACGCTTCTTCAAG	4036
Qy	901	ACCTGCGCGCCGAGCAGACACCCAGAGAGTGAAGACTGAATGACCGACACTGCTCTG	960
Db	4037	ACCTGCGCGCCGAGCAGACACCCAGAGAGTGAAGACTGAATGACCGACACTGCTCTG	4096
Qy	961	GTGCAGAACCGCAACCCCGACTGCAAGACCATCTCTGGCGCTCTCGGCCCCGGCGCAC	1020
Db	4097	GTGCAGAACCGCAACCCCGACTGCAAGACCATCTCTGGCGCTCTCGGCCCCGGCGCAC	4156
Qy	1021	CTGAGAGGATGATGACCGCTGCAAGGGGTGGGGGGCCCCCAGCCACAAGGGCCGGTG	1080
Db	4157	CTGAGAGGATGATGACCGCTGCAAGGGGTGGGGGGCCCCCAGCCACAAGGGCCGGTG	4216
Qy	1081	CTGGCCGAGGCGATGAGCCAGGCCAACCAACGCGTGTATGTACGAAGAGCACTTCAAG	1140
Db	4217	CTGGCCGAGGCGATGAGCCAGGCCAACCAACGCGTGTATGTACGAAGAGCACTTCAAG	4276
Qy	1141	GGCCCCGGGCGATGTCATAGTCTTCAACTGCGGCAAGAGGGCCATCGCCGCAAC	1200
Db	4277	GGCCCCGGGCGATGTCATAGTCTTCAACTGCGGCAAGAGGGCCATCGCCGCAAC	4336
Qy	1201	TGCGGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGGCAAGAGGGCCACGATGAAG	1260
Db	4337	TGCGGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGGCAAGAGGGCCACGATGAAG	4396
Qy	1261	GACTGCACCGAGCGCCAGCCAACTTCTGGGCAAGTCTGGGCCACAGCGCCAGAGGCGCG	1320
Db	4397	GACTGCACCGAGCGCCAGCCAACTTCTGGGCAAGTCTGGGCCACAGCGCCAGAGGCGCG	4456
Qy	1321	CCGGCAACTTCTCTGAGAGCGCGCCCCAGCCACCGCCCCCGCCGAGAGCTTCCCG	1380
Db	4457	CCGGCAACTTCTCTGAGAGCGCGCCCCAGCCACCGCCCCCGCCGAGAGCTTCCCG	4516
Qy	1381	TTTCAGAGAGACACACCCCGGCGCAGAGCAGAGACAGACCGCGACACCTTGACACAG	1440
Db	4517	TTTCAGAGAGACACACCCCGGCGCAGAGCAGAGACAGACCGCGACACCTTGACACAG	4576
Qy	1441	CTGAAGAGCCTGTTCGCAACGACCCCTGAGCCAGTAA	1479
Db	4577	CTGAAGAGCCTGTTCGCAACGACCCCTGAGCCAGTAA	4615
RESULT 10			
US-10-190-435-38			
Sequence 38, Application US/10190435			
Publication No. us20030143248A1			
GENERAL INFORMATION:			
APPLICANT: ZUR MEGEDE, Jan			
APPLICANT: BARNETT, Susan W.			
APPLICANT: LIAN, Ying			
APPLICANT: ENGELBRECHT, Susan			
APPLICANT: VAN RENSBURG, Estrelita J.			
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C			
FILE REFERENCE: P18133_003 / 2302-18133			
CURRENT APPLICATION NUMBER: US/10/190,435			
CURRENT FILING DATE: 2002-12-30			
NUMBER OF SEQ ID NOS: 319			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 38			
LENGTH: 4702			
TYPE: DNA			



```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TVI-gagmod.BM965
US-10-190-435-38

```

Query Match 99.9%; Score 1477.4; DB 15; Length 4702;

Best Local Similarity 99.9%; Pred. No. 9.6e-308;

Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGGGGCCCCGGCCAGCATCTCTGGCGGCGGCACTGAGCGCTTGGAGGCGCATCCGC 60
DB 3224 ATGGGGGCCCCGGCCAGCATCTCTGGCGGCGGCACTGAGCGCTTGGAGGCGCATCCGC 3283
QY 61 CTGGCGCCCGGCGCAAGAGTGTCTCATGATGTAACACCTGTGTGTGGCCAGCCGCGAG 120
DB 3284 CTGGCGCCCGGCGCAAGAGTGTCTCATGATGTAACACCTGTGTGTGGCCAGCCGCGAG 3343
QY 121 CTGGAGAAATTGCGCTTGAACCCCGGCTGTGGAGACCAAGCGAGGGCTTGAAGCATTC 180
DB 3344 CTGGAGAAATTGCGCTTGAACCCCGGCTGTGGAGACCAAGCGAGGGCTTGAAGCATTC 3403
QY 181 ATCCGCGCACTGACCCCGCCCTTGCAGACCCGCGAGGAGAGGTGTAAGCTTGTCAAC 240
DB 3404 ATCCGCGCACTGACCCCGCCCTTGCAGACCCGCGAGGAGAGGTGTAAGCTTGTCAAC 3463
QY 241 ACCGTGGCCACCCTGTACTGCTGTGCAAGAGATGAGGTCCGCGACCAAGAGAGCC 300
DB 3464 ACCGTGGCCACCCTGTACTGCTGTGCAAGAGATGAGGTCCGCGACCAAGAGAGCC 3523
QY 301 CTGGACAAAGATCGAGAGAGAGCAAGCAAGTGCAGACAGAGATCCAGAGGCCGAGGCC 360
DB 3524 CTGGACAAAGATCGAGAGAGAGCAAGCAAGTGCAGACAGAGATCCAGAGGCCGAGGCC 3583
QY 361 GCCGCAAGAGGCAAGGTGAGCCAGAACTACCCCATGTGTGCAAGAACTTGAAGGCCAGATG 420
DB 3584 GCCGCAAGAGGCAAGGTGAGCCAGAACTACCCCATGTGTGCAAGAACTTGAAGGCCAGATG 3643
QY 421 GTGCAACGAGGCATCAAGCCCGCGCACTTGAACGCTTGGGTGAGGTGATCGAGGAGAG 480
DB 3644 GTGCAACGAGGCATCAAGCCCGCGCACTTGAACGCTTGGGTGAGGTGATCGAGGAGAG 3703
QY 481 GCGTTTCAAGCCCGGAGGTGATCCCATGTTTCAACCGCCCTTGAAGCGAGGCCACCCGCCAG 540
DB 3704 GCGTTTCAAGCCCGGAGGTGATCCCATGTTTCAACCGCCCTTGAAGCGAGGCCACCCGCCAG 3763
QY 541 GACCTGAACACGATGTTGAACACCGTGGGCGGCGACACAGCCGCCCATGTGAGATGCTGAAG 600
DB 3764 GACCTGAACACGATGTTGAACACCGTGGGCGGCGACACAGCCGCCCATGTGAGATGCTGAAG 3823
QY 601 GACACCATCAAGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCAGCCGCGGCC 660
DB 3824 GACACCATCAAGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCAGCCGCGGCC 3883
QY 661 ATGCCCCCGGCGCAAGTGCAGAGCCCGCGGAGGAGACATGCGCGGACCAACCAAGCACC 720
DB 3884 ATGCCCCCGGCGCAAGTGCAGAGCCCGCGGAGGAGACATGCGCGGACCAACCAAGCACC 3943
QY 721 CTGGAGGAGACGATCCCTCTGATGACACGACCAACCCGCCCATCCCGTGGGCGGACATTC 780
DB 3944 CTGGAGGAGACGATCCCTCTGATGACACGACCAACCCGCCCATCCCGTGGGCGGACATTC 4003
QY 781 AAGCGGTGATCATCTTGGGCGCTGAACAGATGTGCGGATGTGACAGCCCGGAGCAGTC 840
DB 4004 AAGCGGTGATCATCTTGGGCGCTGAACAGATGTGCGGATGTGACAGCCCGGAGCAGTC 4063
QY 841 CTGGACATCAAGAGGCGCCCAAGAGCCCTTCCGGAATCACTGTGACCGCTTCTTCAAG 900
DB 4064 CTGGACATCAAGAGGCGCCCAAGAGCCCTTCCGGAATCACTGTGACCGCTTCTTCAAG 4123
QY 901 ACCCTGGCGCGGAGAGACACCCGAGAGGTGAAGAACTGGATGACCGACCTTGTCTG 960
DB 4124 ACCCTGGCGCGGAGAGACACCCGAGAGGTGAAGAACTGGATGACCGACCTTGTCTG 4183

```

```

QY 961 GTGCAAGAGCCCAACCCCGAGTGCAGACCATCTCTGCGGCTTGGCGCCCGGCGGCGAGC 1020
DB 4184 GTGCAAGAGCCCAACCCCGAGTGCAGACCATCTCTGCGGCTTGGCGCCCGGCGGCGAGC 4243
QY 1021 CTGGAGAGATGATGATGACCGCTTGCAGAGGCGGTGGCGGCGCCCAAGCACAAGGCCGCGTG 1080
DB 4244 CTGGAGAGATGATGATGACCGCTTGCAGAGGCGGTGGCGGCGCCCAAGCACAAGGCCGCGTG 4303
QY 1081 CTGGCGGAGGAGATGATGACGCGCAACACAGCGGTGATGATGAGCAAGAGCAACTTCAAG 1140
DB 4304 CTGGCGGAGGAGATGATGACGCGCAACACAGCGGTGATGATGAGCAAGAGCAACTTCAAG 4363
QY 1141 GCGCCCGGCGCATGTCATGCTTCAACTGCGGCAAGAGAGGCGCATGCGCGCAAC 1200
DB 4364 GCGCCCGGCGCATGTCATGCTTCAACTGCGGCAAGAGAGGCGCATGCGCGCAAC 4423
QY 1201 TGCCCGCGCCCGGCAAGAGGCGTGTGAAAGTGTGGCAAGAGAGGCGCACCATGTAAG 1260
DB 4424 TGCCCGCGCCCGGCAAGAGGCGTGTGAAAGTGTGGCAAGAGAGGCGCACCATGTAAG 4483
QY 1261 GACTGCAACGAGCGCGCAAGCCCACTTCTGGGCAAGATCTGGCCCAAGCACAAGGCGCGC 1320
DB 4484 GACTGCAACGAGCGCGCAAGCCCACTTCTGGGCAAGATCTGGCCCAAGCACAAGGCGCGC 4543
QY 1321 CCGGCAACTTCTCTGACAGCGCGCCGAGGCCACCGCCCGCGCGAGAGCTTCCGC 1380
DB 4544 CCGGCAACTTCTCTGACAGCGCGCCGAGGCCACCGCCCGCGCGAGAGCTTCCGC 4603
QY 1381 TTGAGAGAGACCAACCCCGGCGCAAGAGCAAGAGCAAGAGCCGAGAGACCTTGAACAGC 1440
DB 4604 TTGAGAGAGACCAACCCCGGCGCAAGAGCAAGAGCAAGAGCCGAGAGACCTTGAACAGC 4663
QY 1441 CTGAAGAGCTGTTGGCAAGACCCCTGAGCCAGTAA 1479
DB 4664 CTGAAGAGCTGTTGGCAAGACCCCTGAGCCAGTAA 4702

RESULT 11
US-10-190-435-17
; Sequence 17, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 4716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagProteinArtmutatRevnef_C
US-10-190-435-17

Query Match 99.9%; Score 1477.4; DB 15; Length 4716;
Best Local Similarity 99.9%; Pred. No. 9.6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGGCGGAGCATCTTGGCGGCGGCAAGTGTGAGCGCTTGGAGGCGCATCCGC 60
DB 7 ATGGGCGCGCGGCGGAGCATCTTGGCGGCGGCAAGTGTGAGCGCTTGGAGGCGCATCCGC 66
QY 61 CTGGCGCCCGGCGCAAGAGTGTCTCATGATGTAAGCACTGTGTGTGGCCAGCGCGGAG 120

```



Db 67 CTGCGCCCCGGGCAAGAGTCTCATGATGAAACCTGTGTGGCCAGCCCGAG 126  
Qy 121 CTGGAAGATTGCTGCTGAAACCCCGGCTGTGTGAGACGACGAGGGCTGCAAGCATC 180  
Db 127 CTGGAAGATTGCTGAAACCCCGGCTGTGTGAGACGACGAGGGCTGCAAGCATC 186  
Qy 181 ATCCGCACTGCAACCCCGGCTGTGAGACGAGGAGAGAGTGAAGGCTGTTCAC 240  
Db 187 ATCCGCACTGCAACCCCGGCTGTGAGACGAGGAGAGAGTGAAGGCTGTTCAC 246  
Qy 241 ACCGTGGACCCCTGACTGCTGCTGAGACGAGAGATGAGTCCGAGCAACGAGAGGC 300  
Db 247 ACCGTGGACCCCTGACTGCTGCTGAGACGAGAGATGAGTCCGAGCAACGAGAGGC 306  
Qy 301 CTGGAAGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Db 307 CTGGAAGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366  
Qy 361 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTGTCAGAACTTGCAGGGCCAGATG 420  
Db 367 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTGTCAGAACTTGCAGGGCCAGATG 426  
Qy 421 GTGCAACAGGCCATCAGCCCGGCACTGGAACGCTGGGTGAAGGTATCGAGAGAG 480  
Db 427 GTGCAACAGGCCATCAGCCCGGCACTGGAACGCTGGGTGAAGGTATCGAGAGAG 486  
Qy 481 GCTTCAGCCCGAGAGGTGATCCCATGTTACCGCCCTTGAAGGAGGCGCCACCCCGAG 540  
Db 487 GCTTCAGCCCGAGAGGTGATCCCATGTTACCGCCCTTGAAGGAGGCGCCACCCCGAG 546  
Qy 541 GACCTGAACAGATGTTGAACACCGTGGGCGGCACACGAGCGGCACTGCAATGCTGAAG 600  
Db 547 GACCTGAACAGATGTTGAACACCGTGGGCGGCACACGAGCGGCACTGCAATGCTGAAG 606  
Qy 601 GACACCATCAAGAGAGGCGCGGAGTGGGAGCCGCGTGCACCCGCTGACGCGCGGCCCC 660  
Db 607 GACACCATCAAGAGAGGCGCGGAGTGGGAGCCGCGTGCACCCGCTGACGCGCGGCCCC 666  
Qy 661 ATCGCCCCCGGCGAGATGCGCGAGCCCGCGGAGGAGATGCGCGGCAACCAACGACCC 720  
Db 667 ATCGCCCCCGGCGAGATGCGCGAGCCCGCGGAGGAGATGCGCGGCAACCAACGACCC 726  
Qy 721 CTGCAAGAGAGATGCTGCTGATGACCAAGAACCCCGCATCCCGTGGGCGCATCTAC 780  
Db 727 CTGCAAGAGAGATGCTGCTGATGACCAAGAACCCCGCATCCCGTGGGCGCATCTAC 786  
Qy 781 AAGCGGTGATCATCTGGGCGCTGAACAGATGCTGCGATGTAACGAGCCCGTGAAGCATC 840  
Db 787 AAGCGGTGATCATCTGGGCGCTGAACAGATGCTGCGATGTAACGAGCCCGTGAAGCATC 846  
Qy 841 CTGGAATCAAGAGAGGCGCCCAAGAGAGCCCTTCGCGCATAGCTGAGACCGCTTCTTCAAG 900  
Db 847 CTGGAATCAAGAGAGGCGCCCAAGAGAGCCCTTCGCGCATAGCTGAGACCGCTTCTTCAAG 906  
Qy 901 ACCCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
Db 907 ACCCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966  
Qy 961 GTGGAAGAGGCAACCCCGCATGCAAGACATCTGCGGCGCTTCGCGCGCGCGCGAGC 1020  
Db 967 GTGGAAGAGGCAACCCCGCATGCAAGACATCTGCGGCGCTTCGCGCGCGCGCGAGC 1026  
Qy 1021 CTGGAAGAGATGATGACCGCTGCGAGGCGGTGGGCGGCGCCAGCAAGAGCGCGCTG 1080  
Db 1027 CTGGAAGAGATGATGACCGCTGCGAGGCGGTGGGCGGCGCCAGCAAGAGCGCGCTG 1086  
Qy 1081 CTGGCGGAGGCGATGAGCGCAACCAACAGCTGATGATGAGAGAGAGAGAGAGAGAG 1140  
Db 1087 CTGGCGGAGGCGATGAGCGCAACCAACAGCTGATGATGAGAGAGAGAGAGAGAGAG 1146  
Qy 1141 GGGCCCCGGGCGATGCTCAAGTGTTCATCTGCGGAGAGAGAGGCGCACTCGCGCGCAAC 1200

Db 1147 GGGCCCCGGGCGATGTCAGAGTCTTCAATCTGGGCAAGAGAGGCGCACATGCCCGCAAC 1206  
Qy 1201 TGCGCGCGCGCGCGAGAGAGGCTGTGGAAGTGGCGCAAGAGAGGCGCACCAAGATGAAG 1260  
Db 1207 TGCGCGCGCGCGCGAGAGAGGCTGTGGAAGTGGCGCAAGAGAGGCGCACCAAGATGAAG 1266  
Qy 1261 GACTGCAACGAGAGCGGCAAGCTTCTTGGGCAAGATCTGGGCGCAAGAGAGAGGCGCGC 1320  
Db 1267 GACTGCAACGAGAGCGGCAAGCTTCTTGGGCAAGATCTGGGCGCAAGAGAGAGGCGCGC 1326  
Qy 1321 CCCGCAACTTCTCTCAGAGCGCGCCGAGCCACCGCGCGCGCGCGAGAGCTTCCGC 1380  
Db 1327 CCCGCAACTTCTCTCAGAGCGCGCCGAGCCACCGCGCGCGCGCGAGAGCTTCCGC 1386  
Qy 1381 TTGAGAGAGACCAACCCCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Db 1387 TTGAGAGAGACCAACCCCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446  
Qy 1441 CTGAAGAGCTGTGTCGCAACGACCCCTGAGCCAGTAA 1479  
Db 1447 CTGAAGAGCTGTGTCGCAACGACCCCTGAGCCAGAAA 1485

RESULT 12  
US-10-190-305A-13  
; Sequence 13, Application US/10190305A  
; Publication No. US20030198621A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARRETT, Susan  
; APPLICANT: LIAN, Ying  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR  
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USBS THERIOF  
; FILE REFERENCE: 2302-18702 / 18702.002  
; CURRENT APPLICATION NUMBER: US/10/190,305A  
; CURRENT FILING DATE: 2002-07-05  
; NUMBER OF SEQ ID NOS: 93  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 4716  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: GasProtInaRtmutrRevNeI\_C  
US-10-190-305A-13

Query Match 99.9%; Score 1477.4; DB 16; Length 4716;  
Best Local Similarity 99.9%; Pred. No. 9,6e-308; Indels 0; Gaps 0;  
Matches 1478; Conservative 0; Mismatches 1;  
Qy 1 ATGGGCGCGCGCGAGCATCTGCGGCGGCAAGTGAAGCGCTGGAGGCGCATCCGC 60  
Db 7 ATGGGCGCGCGCGAGCATCTGCGGCGGCAAGTGAAGCGCTGGAGGCGCATCCGC 66  
Qy 61 CTGGCGCGCGCGGCAAGAGTGTCAATGATGAAGCACTGTGTGTGGCGGAGCGCGAG 120  
Db 67 CTGGCGCGCGCGGCAAGAGTGTCAATGATGAAGCACTGTGTGTGGCGGAGCGCGAG 126  
Qy 121 CTGGAAGATTGCTGCTGAAACCCCGGCTGTGTGAGACGACGAGGGCTGCAAGCATC 180  
Db 127 CTGGAAGATTGCTGCTGAAACCCCGGCTGTGTGAGACGACGAGGGCTGCAAGCATC 186  
Qy 181 ATCCGCACTGCAACCCCGGCTGTGAGACGAGGAGAGAGTGAAGGCTGTTCAC 240  
Db 187 ATCCGCACTGCAACCCCGGCTGTGAGACGAGGAGAGAGTGAAGGCTGTTCAC 246  
Qy 241 ACCGTGGACCCCTGACTGCTGCTGAGACGAGAGATGAGTCCGAGCAACGAGAGGC 300  
Db 247 ACCGTGGACCCCTGACTGCTGCTGAGACGAGAGATGAGTCCGAGCAACGAGAGGC 306  
Qy 301 CTGGAAGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

```
Db 307 CTGACCAAGATCGAGAGGAGAGCAAACTGTCACAGAGATCCAGAGCCGAGGCC 366
QY 361 GCGGCAAGAGGAGAGTGAAGCAAGAACTATCCCATGTGTCAGAACTTGAGAGGCCAGAG 420
Db 367 GCGGCAAGAGGAGAGTGAAGCAAGAACTATCCCATGTGTCAGAACTTGAGAGGCCAGAG 426
QY 421 GTGCAACAGAGGAGTGAAGCCCGGCAACCTGAAACGCTGGGTGAAGGTGATGAGAGAA 480
Db 427 GTGCAACAGAGGAGTGAAGCCCGGCAACCTGAAACGCTGGGTGAAGGTGATGAGAGAA 486
QY 481 GCGTTACAGCCCGGAGGTGATCCCATGTTTCAACCGCTTGAAGAGAGGCCCGCCAG 540
Db 487 GCGTTACAGCCCGGAGGTGATCCCATGTTTCAACCGCTTGAAGAGAGGCCCGCCAG 546
QY 541 GACTGAAACAGATGTTGAACAACCGGAGGAGGCAACAGAGGCCGATGAGATGCTGAAG 600
Db 547 GACTGAAACAGATGTTGAACAACCGGAGGAGGCAACAGAGGCCGATGAGATGCTGAAG 606
QY 601 GACACCATCAAGAGAGGAGCCGCGAGTGGACCGCGTGACACCCGCTGACCGCGGCC 660
Db 607 GACACCATCAAGAGAGGAGCCGCGAGTGGACCGCGTGACACCCGCTGACCGCGGCC 666
QY 661 ATCGCCCCCGGCGCAATGCGCGAAGCCCGCGGAGCAAGCATGCGCGGCAACAGCACC 720
Db 667 ATCGCCCCCGGCGCAATGCGCGAAGCCCGCGGAGCAATGCGCGGCAACAGCACC 726
QY 721 CTGCAAGAGCAGATGCGCTGATGATGACAGCAACCCCGCATCCGCTGGGCGACATCTAC 780
Db 727 CTGCAAGAGCAGATGCGCTGATGATGACAGCAACCCCGCATCCGCTGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCGCTGAAACAAAGATGTCGATGATGATGACGCCCGTGAGCATC 840
Db 787 AAGCGGTGATCATCTGGGCGCTGAAACAAAGATGTCGATGATGATGACGCCCGTGAGCATC 846
QY 841 CTGAGCATCAAGAGAGGAGCCCGCAAGAGCCCTTCCGAGCTAGCTGAGACCGCTTTCAAG 900
Db 847 CTGAGCATCAAGAGAGGAGCCCGCAAGAGCCCTTCCGAGCTAGCTGAGACCGCTTTCAAG 906
QY 901 ACCCTGCGGCGGAGAGAGCAACCCGAGAGTGAAGAACTGGATGACGAGACCCCTGCTG 960
Db 907 ACCCTGCGGCGGAGAGAGCAACCCGAGAGTGAAGAACTGGATGACGAGACCCCTGCTG 966
QY 961 GTGCAAAAGCCAAACCCCGACTGCAAGACATCTTGCGGCTCTGCGGCCCGGCGCAGC 1020
Db 967 GTGCAAAAGCCAAACCCCGACTGCAAGACATCTTGCGGCTCTGCGGCCCGGCGCAGC 1026
QY 1021 CTGAGAGAGATGATGACCCGCTGCGAGGCGGTGGGCGGCCCGCAAGAGCCCGCTG 1080
Db 1027 CTGAGAGAGATGATGACCCGCTGCGAGGCGGTGGGCGGCCCGCAAGAGCCCGCTG 1086
QY 1081 CTGAGCGAGGAGTGAAGAGGAGCAACACAGAGTGAATGATGATGATGATGATGATGATG 1140
Db 1087 CTGAGCGAGGAGTGAAGAGGAGCAACACAGAGTGAATGATGATGATGATGATGATGATG 1146
QY 1141 GCGCCCCGCGGAGTCTGCAAGTCTTCAACTGCGGCAAGAGAGGCCCATGCGCCGCAAC 1200
Db 1147 GCGCCCCGCGGAGTCTGCAAGTCTTCAACTGCGGCAAGAGAGGCCCATGCGCCGCAAC 1206
QY 1201 TGCCTGCGGCCCCCGCAAGAGGAGTCTGAAAGTGGGCAAGAGAGGCCCATGATGAAG 1260
Db 1207 TGCCTGCGGCCCCCGCAAGAGGAGTCTGAAAGTGGGCAAGAGAGGCCCATGATGAAG 1266
QY 1261 GACTGCAACAGAGGAGGAGGAGCAACTCTGAGGCAAGATCTGAGCCGAGCAAGAGGCCG 1320
Db 1267 GACTGCAACAGAGGAGGAGGAGCAACTCTGAGGCAAGATCTGAGCCGAGCAAGAGGCCG 1326
QY 1321 CCGGCAACTTCTCTGAGAGGCGCGCCGAGCCCAAGCCCGCCGCGAGAGCTTCCGCG 1380
Db 1327 CCGGCAACTTCTCTGAGAGGCGCGCCGAGCCCAAGCCCGCCGCGAGAGCTTCCGCG 1386
QY 1381 TTTCGAGAGAACCAACCCCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1387 TTTCGAGAGAACCAACCCCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
```

```
QY 1441 CTGAGAGGCTGTTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479
Db 1447 CTGAGAGGCTGTTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485

RESULT 13
US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-435-20

Query Match 99.8%; Score 1475.8; DB 15; Length 2742;
Best Local Similarity 99.9%; Pred. No. 2.2e-307;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGGAGGATCTTGCGCGCGGCGCAAGCTGAGCGCTGGAGCGCATCCGCG 60
Db 7 ATGGGCGCGCGCGGAGGATCTTGCGCGCGGCGCAAGCTGAGCGCTGGAGCGCATCCGCG 66
QY 61 CTGCGCGCGCGCGGAGAAAGTGTCAATGATGAAAGCACTGTGTGTGGCCAGCCGCGAG 120
Db 67 CTGCGCGCGCGGAGAAAGTGTCAATGATGAAAGCACTGTGTGTGGCCAGCCGCGAG 126
QY 121 CTGAGAGATTCGCGCTGAAACCCCGGCTGTGAGAGCAACAGAGGCTGCAAGCAGATC 180
Db 127 CTGAGAGATTCGCGCTGAAACCCCGGCTGTGAGAGCAACAGAGGCTGCAAGCAGATC 186
QY 181 ATCGCGAGCTGCAACCCCGCTGCAAGCCGCGAGCGAGAGAGCTGAAGAGCTTGTCAAC 240
Db 187 ATCGCGAGCTGCAACCCCGCTGCAAGCCGCGAGCGAGAGAGCTGAAGAGCTTGTCAAC 246
QY 241 ACCGTGGCAGCCCTGTACTGTGGTGCAGAGAAAGATCCAGAGTCCCGGCAACCAAGAGGCC 300
Db 247 ACCGTGGCAGCCCTGTACTGTGGTGCAGAGAAAGATCCAGAGTCCCGGCAACCAAGAGGCC 306
QY 301 CTGCAACAAGATGAGAGAGAGAGCAAGAACTGTCAGAGCAAGAGATTCAGAGCGAGGCC 360
Db 307 CTGCAACAAGATGAGAGAGAGAGCAAGAACTGTCAGAGCAAGAGATTCAGAGCGAGGCC 366
QY 361 GCGGCAAGAGGAGAGTGAAGAGGAGCAAGAACTATCTGTCAGAACTTGCAAGAGGCCAGATG 420
Db 367 GCGGCAAGAGGAGAGTGAAGAGGAGCAAGAACTATCTGTCAGAACTTGCAAGAGGCCAGATG 426
QY 421 GTGCAACAGAGGAGTGAAGAGGAGGAGCAAGAACTATCTGTCAGAACTTGCAAGAGGCCAGATG 480
Db 427 GTGCAACAGAGGAGTGAAGAGGAGGAGCAAGAACTATCTGTCAGAACTTGCAAGAGGCCAGATG 486
QY 481 GCGTTACAGCCCGGAGGTGATCCCATGTTTCAACCGCTTGAAGAGAGGCCCGCCAG 540
Db 487 GCGTTACAGCCCGGAGGTGATCCCATGTTTCAACCGCTTGAAGAGAGGCCCGCCAG 546
QY 541 GACTGAAACAGATGTTGAACAACCGTGGGCGGCAACAGAGGCCGATGAGATGCTGAAG 600
```

```

Db      547 GACCTGAACGAGTGTGAACACCGTGGCGGCGCACAGCCGCCATGCAAGTGCAGAG 606
Qy      601 GACACCATCAAGAGAGAGGCGCCGAGTGGACCGGTGTCAACCCGTGACGCGGCGCC 660
Db      607 GACACCATCAAGAGAGAGGCGCCGAGTGGACCGGTGTCAACCCGTGACGCGGCGCC 666
Qy      661 ATGCGCGCGCGAGATGCGGAGCGCCGCGGAGGACATGCGCGGCGCACAGCACCC 720
Db      667 ATGCGCGCGCGAGATGCGGAGCGCCGCGGAGGACATGCGCGGCGCACAGCACCC 726
Qy      721 CTGCAAGAGCAGATGCGCTGTGATGACAGCAACCCGCCCATCCCGTGGGCGACATCTAC 780
Db      727 CTGCAAGAGCAGATGCGCTGTGATGACAGCAACCCGCCCATCCCGTGGGCGACATCTAC 786
Qy      781 AAGCGGTGATCATCTGGGCGCTGAACAAGATGTCGGATGTACAGCCCGCTGAGCATC 840
Db      787 AAGCGGTGATCATCTGGGCGCTGAACAAGATGTCGGATGTACAGCCCGCTGAGCATC 846
Qy      841 CTGGAATCAAGAGAGGCGCCCAAGAGCGCTTCCGGAATAGTGAACCGCTTCTTCAAG 900
Db      847 CTGGAATCAAGAGAGGCGCCCAAGAGCGCTTCCGGAATAGTGAACCGCTTCTTCAAG 906
Qy      901 ACCCTGCGCGCGAGAGAGCAACCAAGAGGTGAAGAACTGGATGACCGACACCTGCTG 960
Db      907 ACCCTGCGCGCGAGAGAGCAACCAAGAGGTGAAGAACTGGATGACCGACACCTGCTG 966
Qy      961 GTGCAAGAGCGCAACCCCGCATGTGCAAGACATCTGCGCGCTTCTGGCGCGCGCAC 1020
Db      967 GTGCAAGAGCGCAACCCCGCATGTGCAAGACATCTGCGCGCTTCTGGCGCGCGCAC 1026
Qy      1021 CTGGAAGATGATGATGACCGCTGCGAGGCGGTGGGCGCGCCAGCAACAAGCGCGCGT 1080
Db      1027 CTGGAAGATGATGATGACCGCTGCGAGGCGGTGGGCGCGCCAGCAACAAGCGCGT 1086
Qy      1081 CTGCGCGAGCGATGAGCGAGCGCAACAACAGCGTATGATGACAGAAGCAACTTCAAG 1140
Db      1087 CTGCGCGAGCGATGAGCGAGCGCAACAACAGCGTATGATGACAGAAGCAACTTCAAG 1146
Qy      1141 GCGCGCGCGCGCATGTGCAAGTCTTCAACTGCGCGAGAGAGGCGCAACTGCGCGCAC 1200
Db      1147 GCGCGCGCGCGCATGTGCAAGTCTTCAACTGCGCGAGAGAGGCGCAACTGCGCGCAC 1206
Qy      1201 TGGCGCGCGCGCGCAAGAGGCGTGTGGAAGTCCGCGAGAGAGGCGCAACAGATGAAG 1260
Db      1207 TGGCGCGCGCGCGCAAGAGGCGTGTGGAAGTCCGCGAGAGAGGCGCAACAGATGAAG 1266
Qy      1261 GACTGCAACGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCGAGCGCAACAAGGCGCG 1320
Db      1267 GACTGCAACGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCGAGCGCAACAAGGCGCG 1326
Qy      1321 CCGCGCAACTTCTGCAAGCGCGCGCGAGCGCCACCGCGCGCGCGCGAGAGCTTCCGC 1380
Db      1327 CCGCGCAACTTCTGCAAGCGCGCGCGAGCGCCACCGCGCGCGCGCGAGAGCTTCCGC 1386
Qy      1381 TTGCGAGAGACACACCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db      1387 TTGCGAGAGACACACCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
Qy      1441 CTGAAGAGCTGTTGCGCAAGACCCCTGAGCCAGTAA 1479
Db      1447 CTGAAGAGCTGTTGCGCAAGACCCCTGAGCCAGTAA 1485

```

```

RESULT 14
; Sequence 15, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

```

```

; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-305A-15

Query Match      99.8%; Score 1475.8; DB 16; Length 2742;
Best Local Similarity 99.9%; Pred. No. 2.2e-307;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ATGGGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGGAAGCTGGAGCGCATTCGC 60
Db      7  ATGGGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGGAAGCTGGAGCGCATTCGC 66
Qy      61  CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCGAGCGCGAG 120
Db      67  CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCGAGCGCGAG 126
Qy      121  CTGGAAGAGTGTGCTGAACCCCGGCTGTGTGAAGACAGCGAGGCTGCAAGCATC 180
Db      127  CTGGAAGAGTGTGCTGAACCCCGGCTGTGTGAAGACAGCGAGGCTGCAAGCATC 186
Qy      181  ATCCGCGAGCTGCACCCCGCGCTGCAGACCGGCGAGGAGCTGAAGAGCTGTTCAC 240
Db      187  ATCCGCGAGCTGCACCCCGCGCTGCAGACCGGCGAGGAGCTGAAGAGCTGTTCAC 246
Qy      241  ACCGTGCGCACCTGTACTGTGCTGCACGAGAAAGTCAAGTCCGCGACCAAGAGGCG 300
Db      247  ACCGTGCGCACCTGTACTGTGCTGCAGAGAAAGTCAAGTCCGCGACCAAGAGGCG 306
Qy      301  CTGGAACAGATGAGAGAGAGAGCAAGCAAGTCCAGCAAGAAAGATTCAGAGGCGCG 360
Db      307  CTGGAACAGATGAGAGAGAGAGCAAGCAAGTCCAGCAAGAAAGATTCAGAGGCGCG 366
Qy      361  GCCGCAAGAGGCGAGGTGAGCGAGCAACCCCATCTGTGCGAAGCTTCGAGGCGCAG 420
Db      367  GCCGCAAGAGGCGAGGTGAGCGAGCAACCCCATCTGTGCGAAGCTTCGAGGCGCAG 426
Qy      421  GTGCAACGAGCGCATCAGCGCGCGCAACCTTGAACCTTGGGTGAAGTATGAGAGAG 480
Db      427  GTGCAACGAGCGCATCAGCGCGCGCAACCTTGAACCTTGGGTGAAGTATGAGAGAG 486
Qy      481  GCTTTAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCCACCCGAG 540
Db      487  GCTTTAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCCACCCGAG 546
Qy      541  GACTGAACACAGATGTTGAACAACCTGTGGCGGCGCAACAGGCGCGCATATCTGAAG 600
Db      547  GACTGAACACAGATGTTGAACAACCTGTGGCGGCGCAACAGGCGCGCATATCTGAAG 606
Qy      601  GACACCATCAAGAGAGAGGCGCGAGGTGGAGCGCGGTGCAACCCCGTGCACGCGGCGCC 660
Db      607  GACACCATCAAGAGAGAGGCGCGAGGTGGAGCGCGGTGCAACCCCGTGCACGCGGCGCC 666
Qy      661  ATGCGCGCGCGAGATGCGGAGCGCGCGGCGAGCGACATGCGCGGCGACCGACGACCC 720
Db      667  ATGCGCGCGCGAGATGCGGAGCGCGCGGCGAGCGACATGCGCGGCGACCGACGACCC 726
Qy      721  CTGCAAGAGCAGATGCGCTGTGATGACAGCAACCCGCCCATCCCGTGGGCGACATCTAC 780
Db      727  CTGCAAGAGCAGATGCGCTGTGATGACAGCAACCCGCCCATCCCGTGGGCGACATCTAC 786
Qy      781  AAGCGGTGATCATCTGGGCGCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 840
Db      787  AAGCGGTGATCATCTGGGCGCTGAACAAGATCTGCGGATGTACAGCCCGCTGAGCATC 846

```

QY 841 CTGACATCAAGCAGAGGCCCCCAAGAGCCCTTCCGCACTAGTGAGCCGCTTTCTTCAAG 900  
DB 847 CTGACATCAAGCAGAGGCCCCCAAGAGCCCTTCCGCACTAGTGAGCCGCTTTCTTCAAG 906  
QY 901 ACCCTGCGCGCGAGAGAGCAACCCAGAGGTGAAGAACTGGATGACCGACACCTTGCTG 960  
DB 907 ACCCTGCGCGCGAGAGAGCAACCCAGAGGTGAAGAACTGGATGACCGACACCTTGCTG 966  
QY 961 GTGAGAAACGCAACCCCGACCTGCAAGAACATCTGCGCGCTCTGGGCCCCGCGCGAGC 1020  
DB 967 GTGAGAAACGCAACCCCGACCTGCAAGAACATCTGCGCGCTCTGGGCCCCGCGCGAGC 1026  
QY 1021 CTGAGAGATGATGATGACCGCTTCCAGAGGCGTGGCGGCCCCAGCACAGGCCCCGCTG 1080  
DB 1027 CTGAGAGATGATGATGACCGCTTCCAGAGGCGTGGCGGCCCCAGCACAGGCCCCGCTG 1086  
QY 1081 CTGCGCGAGGCGATGACCGAGCCCAACACAGCGTGATGATGACAGAGCAACTTTCAG 1140  
DB 1087 CTGCGCGAGGCGATGACCGAGCCCAACACAGCGTGATGATGACAGAGCAACTTTCAG 1146  
QY 1141 GCGCGCGCGCGATGCTCAAGTGTTCAGCTGGCGCAAGAGAGGCGCACTGCGCGCGCAAC 1200  
DB 1147 GCGCGCGCGCGATGCTCAAGTGTTCAGCTGGCGCAAGAGAGGCGCACTGCGCGCGCAAC 1206  
QY 1201 TGCGCGCGCGCGCGAGAGAGGCTGTGGAAGTGGCGCGCAAGAGAGGCGCACTGAG 1260  
DB 1207 TGCGCGCGCGCGCGAGAGAGGCTGTGGAAGTGGCGCGCAAGAGAGGCGCACTGAG 1266  
QY 1261 GACTGACCGAGAGCGCGAGCCCACTTCTGGCGAAGATTTGGCCCAAGCCCAAGAGCGCGC 1320  
DB 1267 GACTGACCGAGAGCGCGAGCCCACTTCTGGCGAAGATTTGGCCCAAGAGCGCGC 1326  
QY 1321 CCGCGCACTTCTGAGAGCGCGCGCGAGCCCAAGCGCGCGCGCGAGACTTTCGCG 1380  
DB 1327 CCGCGCACTTCTGAGAGCGCGCGCGAGCCCAAGCGCGCGCGCGAGACTTTCGCG 1386  
QY 1381 TTGAGAGAGACCAACCCCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
DB 1387 TTGAGAGAGACCAACCCCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446  
QY 1441 CTGAAGAGCTGTGTCGCAAGCAACCCCTGAGCCAGTAA 1479  
DB 1447 CTGAAGAGCTGTGTCGCAAGCAACCCCTGAGCCAGTAA 1485

RESULT 15  
US-10-190-435-9  
; Sequence 9, Application US/10190435  
; Publication No. US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGELBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Estrelita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: PRI9133.003 / 2302-18133  
; CURRENT APPLICATION NUMBER: US/10/190,435  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 3930  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GagComp1Polmut\_C  
US-10-190-435-9

Query Match 99.8%; Score 1475.8; DB 15; Length 3930;  
Best Local Similarity 99.9%; Pred. No. 2.1e-307;  
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCGGC 60  
DB 7 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCGGC 66  
QY 61 CTGCGCGCGCGCGCAAGAACTGTCTACATGATGAAGCACTGTGTGTGGCCAGCCGAG 120  
DB 67 CTGCGCGCGCGCGCAAGAACTGTCTACATGATGAAGCACTGTGTGTGGCCAGCCGAG 126  
QY 121 CTGAGAAAGTTGCGCTGGAACCCCGCGCTGTGAGAGCAAGAGGCGTGAACAGATC 180  
DB 127 CTGAGAAAGTTGCGCTGGAACCCCGCGCTGTGAGAGCAAGAGGCGTGAACAGATC 186  
QY 181 ATCCGCGAGCTGCAACCCCGCTGCAAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 187 ATCCGCGAGCTGCAACCCCGCTGCAAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAG 246  
QY 241 ACCGTGCGCACCTGTGATCTGCTGCAAGAGAGATCGAGGTCCCGACACCAAGAGGCGC 300  
DB 247 ACCGTGCGCACCTGTGATCTGCTGCAAGAGAGATCGAGGTCCCGACACCAAGAGGCGC 306  
QY 301 CTGAGCAAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 307 CTGAGCAAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366  
QY 361 GCGGCAAGAGGCGAGGTGAGCGAGAACTAACCTGCGAGAACTTGCAGAGGCGCAGATG 420  
DB 367 GCGGCAAGAGGCGAGGTGAGCGAGAACTAACCTGCGAGAACTTGCAGAGGCGCAGATG 426  
QY 421 GTGACACAGGCTCATAGCCCGCGCAACCTTGAAAGCTGTGAGAGTATGAGAGAGAG 480  
DB 427 GTGACACAGGCTCATAGCCCGCGCAACCTTGAAAGCTGTGAGAGTATGAGAGAGAG 486  
QY 481 GCTTCAAGCGCGCGAGTATCCCATGTTCAACCGCGCTGAGCGAGGCGGCAACCCCGAG 540  
DB 487 GCTTCAAGCGCGCGAGTATCCCATGTTCAACCGCGCTGAGCGAGGCGGCAACCCCGAG 546  
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGCGGCGCAAGCGCGCATATGAGCTGAG 600  
DB 547 GACCTGAACACGATGTTGAACACCGTGGCGCGGCGCAAGCGCGCATATGAGCTGAG 606  
QY 601 GACACATCAACGAGAGAGCGCGCGAGTGGAGACCGCGTGACACCCCGTGACAGCGCGGCC 660  
DB 607 GACACATCAACGAGAGAGCGCGCGAGTGGAGACCGCGTGACACCCCGTGACAGCGCGGCC 666  
QY 661 ATCGCGCGCGCGAGTGCAGAGCGCGCGCGAGCGGCAATCGCGCGGCAACCAAGAGAGC 720  
DB 667 ATCGCGCGCGCGAGTGCAGAGCGCGCGCGAGCGGCAATCGCGCGGCAACCAAGAGAGC 726  
QY 721 CTGAGAGAGAGATGCGCTGATGACAGCAACCCCGCATCCCGTGGAGCGATCTAC 780  
DB 727 CTGAGAGAGAGATGCGCTGATGACAGCAACCCCGCATCCCGTGGAGCGATCTAC 786  
QY 781 AAGCGGTGATCATCTGGGCTTGAACAAATCGTGGAGATGACAGCCCGTGAGCATC 840  
DB 787 AAGCGGTGATCATCTGGGCTTGAACAAATCGTGGAGATGACAGCCCGTGAGCATC 846  
QY 841 CTGACATCAAGCAGAGGCGCGCAAGAGAGCCCTTCCGCGATGAGAGAGAGAGAGAG 900  
DB 847 CTGACATCAAGCAGAGGCGCGCAAGAGAGCCCTTCCGCGATGAGAGAGAGAGAGAG 906  
QY 901 ACCCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
DB 907 ACCCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966  
QY 961 GTGAGAAACGCAACCCCGAGCTGCAAGCAACTCTGCGGCTTGGGCGCGCGCGAGC 1020  
DB 967 GTGAGAAACGCAACCCCGAGCTGCAAGCAACTCTGCGGCTTGGGCGCGCGCGAGC 1026  
QY 1021 CTGAGAGAGATGATGATGACCGCTTCCAGAGGCGTGGCGGCCCCAGCACAGGCGCGCTG 1080  
DB 1027 CTGAGAGAGATGATGATGACCGCTTCCAGAGGCGTGGCGGCCCCAGCACAGGCGCGCTG 1086

QY 1081 CTGCGGAGGCGATGAGCCAGGCCAACACACGCGTGATGATGAGAGAGCAACTTCAAG 1140  
 DB 1087 CTGCGCGAGCGATGAGCCAGGCCAACACACGCGTGATGATGAGAGAGCAACTTCAAG 1146  
 QY 1141 GGCGCGGCGGCGATGCTCAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCGCAAC 1200  
 DB 1147 GGCGCGGCGGCGATGCTCAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCGCAAC 1206  
 QY 1201 TGCCTGCGCGCGCGGCAAGAGGCTGCTGGAAGTGCAGCAAGAGGGCCACATGATGAAG 1260  
 DB 1207 TGCCTGCGCGCGCGGCAAGAGGCTGCTGGAAGTGCAGCAAGAGGGCCACATGATGAAG 1266  
 QY 1261 GACTGACCGAGCGGCGGCAAGTCTGCGGCAAGATCTGCGGCAAGTCTGCGGCAAGGCGCGC 1320  
 DB 1267 GACTGACCGAGCGGCGGCAAGTCTGCGGCAAGATCTGCGGCAAGTCTGCGGCAAGGCGCGC 1326  
 QY 1321 CCGGCGCACTTCTCTGAGAGCCGCGCGGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380  
 DB 1327 CCGGCGCACTTCTCTGAGAGCCGCGCGGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGC 1386  
 QY 1381 TTGAGAGAGACCAACCGCGCGCGCGAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 1440  
 DB 1387 TTGAGAGAGACCAACCGCGCGCGCGAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 1446  
 QY 1441 CTGAAGAGCGCTGTTGCGCAAGCAAGCGCGCTGAGCCAGTAA 1479  
 DB 1447 CTGAAGAGCGCTGTTGCGCAAGCAAGCGCGCTGAGCCAGTAA 1485

Search completed: September 1, 2005, 23:17:17  
 Job time : 920.198 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 1, 2005, 21:17:23 | Search time 25.739 Seconds  
(without alignments)  
8578.900 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 2772

Sequence: 1 atgggcgcgcgcgcgcagcat.....acgaccctctgagccagtaa 1479

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US09475704/runtac\_01092005\_160751\_19541/app\_query.fasta\_1.3342  
-DB=Issued Patents\_AA -QFMT=fastaan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09475704.@CEN1\_1.53.@runtac\_01092005\_160751\_19541 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2435	87.8	492	4	US-09-991-258-5
2	2256	81.4	500	2	US-08-816-155B-45
3	2256	81.4	500	3	US-09-079-587-45
4	2249	81.1	500	1	US-08-375-510-1
5	2249	81.1	500	2	US-08-487-657-1
6	2249	81.1	500	4	US-09-309-572-16
7	2249	81.1	500	4	US-09-718-096-16
8	2245	81.1	1350	4	US-09-952-060-35
9	2235	80.6	512	3	US-08-463-210-8
10	2235	80.6	512	3	US-09-124-900-2
11	2235	80.6	512	4	US-08-463-028-8
12	2106	76.0	493	4	US-09-952-060-33

13	2037.5	73.5	478	1	US-08-127-499A-11	Sequence 11, Appl
14	2037.5	73.5	478	1	US-08-482-847-11	Sequence 11, Appl
15	2023	73.0	437	3	US-08-392-794A-2	Sequence 2, Appl1
16	1976	71.3	512	4	US-09-319-588C-4	Sequence 4, Appl1
17	1797.5	64.8	498	1	US-08-470-202-59	Sequence 59, Appl
18	1797.5	64.8	498	1	US-08-471-770-59	Sequence 59, Appl
19	1797.5	64.8	498	2	US-08-468-059-59	Sequence 59, Appl
20	1797.5	64.8	498	3	US-09-109-916-59	Sequence 59, Appl
21	1797.5	64.8	498	3	US-09-886-156-59	Sequence 59, Appl
22	1797.5	64.8	498	4	US-09-886-149-59	Sequence 59, Appl
23	1797.5	64.8	498	4	US-09-886-150-59	Sequence 59, Appl
24	1797.5	64.8	498	4	US-09-886-159-59	Sequence 59, Appl
25	1797.5	64.8	498	4	US-10-326-090-59	Sequence 59, Appl
26	1775.5	64.1	498	1	US-08-470-202-60	Sequence 60, Appl
27	1775.5	64.1	498	1	US-08-471-770-60	Sequence 60, Appl
28	1775.5	64.1	498	2	US-08-468-059-60	Sequence 60, Appl
29	1775.5	64.1	498	3	US-09-109-916-60	Sequence 60, Appl
30	1775.5	64.1	498	4	US-09-886-156-60	Sequence 60, Appl
31	1775.5	64.1	498	4	US-09-886-149-60	Sequence 60, Appl
32	1775.5	64.1	498	4	US-09-886-150-60	Sequence 60, Appl
33	1775.5	64.1	498	4	US-09-886-159-60	Sequence 60, Appl
34	1775.5	64.1	498	4	US-10-326-090-60	Sequence 60, Appl
35	1681	60.6	363	2	US-08-850-049-130	Sequence 130, App
36	1681	60.6	363	2	US-08-050-478-130	Sequence 130, App
37	1681	60.6	363	3	US-09-414-117-130	Sequence 130, App
38	1681	60.6	363	4	US-09-678-437-130	Sequence 130, App
39	1681	60.6	363	4	US-09-943-722-130	Sequence 130, App
40	1536.5	55.4	458	1	US-07-648-796A-5	Sequence 5, Appl1
41	1536.5	55.4	799	1	US-07-648-796A-7	Sequence 7, Appl1
42	1530.5	55.2	337	1	US-07-648-796A-1	Sequence 7, Appl1
43	1514.5	54.6	328	4	US-08-776-188C-77	Sequence 77, Appl
44	1478.5	53.3	518	4	US-09-206-551-45	Sequence 45, Appl1
45	1413	51.0	294	3	US-09-370-368-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-991-258-5  
Sequence 5, Application US/09991258  
Patent No. 6783939  
GENERAL INFORMATION:  
APPLICANT: Olmsted, Robert  
APPLICANT: Keith, Paula  
APPLICANT: Dryga, Sergey  
APPLICANT: Caley, Ian  
APPLICANT: Maughan, Maureen  
APPLICANT: Johnston, Robert  
APPLICANT: Davis, Nancy  
APPLICANT: Swansstrom, Ronald  
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSONES WITH MODIFIED HIV GENES FOR USE  
FILE REFERENCE: 01113.0001U3  
CURRENT FILING DATE: 2001-11-16  
CURRENT FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: 09/902,537  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: 60/216,995  
PRIOR FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 492  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence; No. 6783939 =  
US-09-991-258-5  
Alignment Scores:  
Pred. No.: 2,976-141  
Score: 2435.00  
Length: 492  
Matches: 458

Percent Similarity: 96.96% Conservative: 20  
 Best Local Similarity: 92.90% Mismatches: 13  
 Query Match: 87.84% Indels: 2  
 DB: 4 Gaps: 2

US-09-475-704A-3 (1-1479) x US-09-991-258-5 (1-492)

```

QY 1 ATGGGGCGCCGGCCGAGCATCTCGGCGGGCGGCAAGCTGAGCCGCTGGAGGCGCATCCGC 60
DB 1 MetAlaAlaArgAlaSerIleLeuArgGlyGluIlybLeuAspLysTrpGluLysIleArg 20
QY 61 CTGGCGCCCGCGCGCAAGAAGTGTACATGATGAAGACACCTGATGTGGCGCCAGCGCGAG 120
DB 21 LeuArgProGlyGlybLysbHsIeTyMetLeuLysHsIeValIleTrpAlaSerArgGlu 40
QY 121 CTGGAAGAGTTGGCCCTGAACCCCGGCGCTGTGAGAGACCGAGGCGCTGCAGACAGATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysbLysGlnIle 60
QY 181 ATCCCGCCAGCTGCACCCCGCGCTGCAGACCGGCGAGAGAGCTGAAGCCTGTTCAAC 240
DB 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyArgn 80
QY 241 ACCGTGGCCACCTGTACTGCTGTCAGAGAGATGAGAGTCCGCGACACAGAGAGGCC 300
DB 81 ThrValAlaThrLeuTyrcyValHsIeGluLysIleGluValArgPheThrIyGlnAla 100
QY 301 CTGAGACAGATCGAGAGAGAGAGACAAAGTGCACAGACAGAGATCCAGAGCCAGAGCC 360
DB 101 LeuAspLysIleGluGluGluGlnIleAsnLysCyGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCGGCAAGAGGGAGGTGAGCGCAAGACTACCCCATGTGACAGAACTGAGGGCGCAGATG 420
DB 121 AlaAsp--GlybValValSerGlnAsnTyProIleValGlnbAsnLeuGlnIyGlnMet 139
QY 421 GTGCAACGAGCCATAGCCCGCGCAACCTGAAAGCTGTGAGGTATCGAGAGAGAG 480
DB 140 ValHsIeGlnAlaIleSerProArgThrIleuAsnAlaIleTrpValIleGluGluLys 159
QY 481 GCGTTGAGCCCGAGAGGTGATCCCATGTTCAACCGCCCTGAGAGAGGGCGCCACCCCGAG 540
DB 160 AlaPheSerProGlyValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 179
QY 541 GACCTGAACAGATGTTGAACAACCGTGGCGGCGCACAGCGCGCATGAGATGCTGAAG 600
DB 180 AspLeuAsnThrMetLeuAsnThrValGlyGlnHsIeGlnAlaIleValMetIleMetLeuLys 199
QY 601 GACACCATCAAGAGAGAGCGCCGCGAGTGGAGACCGGCTGACACCCGTCGACCGCGCCCG 660
DB 200 AspThrIleAsnGluGlnAlaIleGluAlaGluTrpArgPheLeuHsIeProValHsIeGlyPro 219
QY 661 ATGCGCCCGCGCGCAAGTGGCGAGCCCGCGCGAGCGACATGATGCGCGGACCAACAGACC 720
DB 220 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGAGAGAGACAGATGCGCTGATGACCAAGACCCCGCATCCCGCTGGCGGACATCTAC 780
DB 240 LeuGlnGlnGlnIleAlaIleThrPheCtnrSerAsnProProIleProValGlyAspIleTy 259
QY 781 AAGCGGTGATCATCTGGGCTGGAACAAGATGTCGAGATTAACAGCCCGCTGAGATC 840
DB 260 LysArgTrpIleIleLeuGlyLeuAsnLysbIleValArgMetLysSerProValSerIle 279
QY 841 CTGGAACATCAAGAGAGGGCCCAAGAGCCCTTCGCGACTACGTGAGACCGCTTCTTCAAG 900
DB 280 LeuAspIleArgGlnIyProLysGluProPheArgAspTyValAspArgPhePheLys 299
QY 901 ACCGTGGCGCGAGAGAGAGACCCAGAGGTGAAGAAATGATGACGACACCCCTGCTG 960
DB 300 ThrLeuArgAlaGluGlnAlaThrGlnGlnValLysAsnTrpPheCtnrAspThrLeuLeu 319
QY 961 GTGCAAGAGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCGCCGCGCGCAAC 1020

```

```

DB 320 ValGlnAsnAlaAsnProAspCysbLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGAGAGAGATGATGACCGCTGCGCAGGGCGCTGGCGGGCGCCGACCAAGAGCCCGCTG 1080
DB 340 LeuGluGluMetMetThrAlaCysGlnIyValGlyGlyProGlyHsIyLysAlaArgVal 359
QY 1081 CTGGCGCGAGATGAGAGAGGACGACAC--AGCGATGATGACAGAAAGCAACTTC 1137
DB 360 LeuAlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe 379
QY 1138 AAGGCGCCCGCGCATGTCAAGTCTTCAACTGCGGCAAGAGAGGCGACATGCGCCGC 1197
DB 380 LysGlyProArgArgIleValLysCysPheAsnCyGlyLysGluGlnHsIeIleAlaArg 399
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCGGCAAGAGGCGCACCATG 1257
DB 400 AsnCysArgAlaProArgLysbLysGlyCysbTrpCysGlyLysGluGlnHsIeMet 419
QY 1258 AAGACTGACACCGAGCGCGAGCGCAACTTCTGGGCAAGATCTGAGCCGACAGAGGCG 1317
DB 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlybIleTrpProSerHsIyGly 439
QY 1318 GCGCCCGCAACTTCTGTCAGAGCGCGCGCGAGCCGACCGCGCGCGAGAGCTTC 1377
DB 440 ArgProGlyAsnPheLeuGlnAlaAsnArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 GCGTTGAGAGAGACCAACCCCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
DB 460 ArgPheGluGluThrThrProAlaProLysGlnIyProIleGluArgGluProLeuThr 479
QY 1438 AGCTGAAGAGCTGTTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1496
DB 480 SerLeuLysSerLeuPheGlySerPheProLeuSerGln 492

RESULT 2
US-08-816-155B-45
; Sequence 45, Application US/08816155B
; Patent No. 5990091
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,155B
; FILING DATE: 12-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KOMALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454310-2990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0500
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:

```





```

1 SOFTWARE: PatentIn Release #1.0, Version #1.3
2
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/09/079,587
5
6 FILING DATE:
7
8 CLASSIFICATION:
9
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US 08/816,155
12 FILING DATE: 12-MAR-1997
13 ATTORNEY/AGENT INFORMATION:
14 NAME: KOMALSKI, THOMAS J.
15
16 REGISTRATION NUMBER: 32,147
17
18 REFERENCE/DOCKET NUMBER: 454310-2990
19
20 TELECOMMUNICATION INFORMATION:
21
22 TELEPHONE: 212-588-0800
23
24 TELEFAX: 212-588-0500
25
26 INFORMATION FOR SEQ ID NO: 45:
27
28 SEQUENCE CHARACTERISTICS:
29
30 LENGTH: 500 amino acids
31
32 TYPE: amino acid
33
34 STRANDEDNESS: n/a
35
36 TOPOLOGY: linear
37
38 MOLECULE TYPE: amino acid
39
40 US-09-079-587-45

```

Alignment Scores:	
Pred. No.:	2,59e-130
Score:	2256.00
Percent Similarity:	92.42%
Best Local Similarity:	84.63%
Query Match:	81.39%
DB:	3
Gaps:	5
length:	5000
Matches:	422
Conservative:	39
Mismatches:	28
Indels:	10
Gaps:	5

US-09-475-704A-3 (1-1479) x US-09-079-587-45 (1-500)

[illegible]

QY	592	ATGCTGAAGGACACACCTCAACAGAGAGGCGCCGAGTGGGACCGGTGACACCCCGTGGAC	651
Db	200	MetLeuNyBgIutnrlIeaNgIngluAlaIaIgutrPaapArGValHIsProValHIs	219
QY	652	GCCGAGCCCATCGCGCCCGGCGCAATATGGCGAGCCCGCGGACGACATCGCGCGACCC	711
Db	220	AlaGlyProIleAlaProGlyGImeCarGeluProArGlySerAspIleAlaGlyThr	239
QY	712	ACGACGACCTTCGAGAGACAGATGCGCTGTGATGACGACAAACCCCCATCCCGTGGCC	771
Db	240	ThSerThrIeaNgIngluInllIeGlyTrpMetThrAenAnProIleProValGly	259
QY	772	GACATCTCAACAGCGGTGTGATCATCTCGGCGCTGAAACAAGATCTGCGGATGTACAGCCCC	831
Db	260	GluIleTyrlAsArgTpIleIleuGlyLeuAenNySllIeValAryMetCysIerPro	279
QY	832	GTGAGCATCTCCGACATCAACAGAGGCGCCCAAGAGACCCCTTCGCGCATCGTGAAGCGC	891
Db	280	ThSerIleuAapIleArGInGlyProIlySgluProPheArGAspTyValAspArg	299
QY	892	TTCTTCAAGACCTTCGCGCGCGCGAGACGACACCCAGAGGTGAGAACTGATGACCGAC	951
Db	300	PheTyrlSerThrIeArGAlaGluInAlaSerGInGluValIyAsnTrpMetThrGlu	319
QY	952	ACCTGTCTGTGTGAGAAACGCCAACCCGCACTGTGCAAGACATCTCGCGGCTCTCGGCCCC	1011
Db	320	ThrIeuIeuValGlnAsnAlaAsnProAspCysIySthrIleIeuAlaIeulIyPro	339
QY	1012	GGCCCGACGCTCGAGAGAGATGATACACGCTCTCGCGAGGCGTGTGGCGGCCCGACCAAG	1072
Db	340	AlaAlaThrIeuGluGImeMetThrAlaCysGlnGlyValGlyGlyProGlyHisIyS	359
QY	1072	GCCCGCGTGTGCGCGAGCGCATGAGCCGACCAACACGACG-----GTGATGTGCA	1122
Db	360	AlaArgValIleuAlaGluAlaMetSerGlnValThrAenSerAlaThrIleMetMetGln	379
QY	1126	AAGAGCAACTTCMAAGGCGCCCGCGCGCATCTGTCMAAGTGTCTCAATCTGGCGCAAGAGGCG	1186
Db	380	ArgGlyAenPheArGAsnGlnArGlyIleValIySvPheAsnCysGlyIySgluGly	399
QY	1186	CACATCGCGCCCACTGCGCGCGCGCCCGCGCAAGAGGCGCTCTGGAATGTCGGCAAGAG	1246
Db	400	HisThrAlaArGAsnCysArGAlaProArGlySlySglyCysTrpIyCysGlyIySglu	419
QY	1246	GGCCACCAAGATGAAGACTGTGACCGAGCGCAAGCCCACTTCTGGGCAAGATCTGGCCC	1306
Db	420	GlyHisGImeTlyAspCysThrGluIyArgGlnAlaAenPheIeuGlyLysIleTrpPro	439
QY	1306	AGCCACMAAGGCGCGCCCGCGCAACTTCTGTGAGAGCGCGCCCGAGACCAACGCGCCCCCCC	1366
Db	440	SerTyrlSgluIyArGProGlyIyAsnPheIeuGlnSerArGProGluProThrIaIProPro	459
QY	1366	GCCGAGACTTCCCG-----TTGAGAGAGACCAACCCCGCGCAGAAAGAGAGAGCAAG	1411
Db	460	GluGluSerPheArGSerGlyValGluIyThrThrIleProProGlnIySgluIyProIle	479
QY	1420	GACCGCGAG-----ACCTGACCAAGCCTGTGAAGAGCCTGTTCGCAACGACCCCGTGAAC	1477
Db	480	AspIySgluIeuTyrlProIeuThrSerIeuArGSerIeuPheGlyIyAsnAspProSerSer	499
QY	1474	CAG 1476	
Db	500	Gln 500	

RESULT 4  
US-08-375-510-1  
; Sequence 1, Application US/08375510  
; Patent No. 5576421  
; GENERAL INFORMATION:  
; APPLICANT: Satco, Atsushi  
; APPLICANT: Singsawa, Hideo  
; APPLICANT: Nakata, Atsuo

TITLE OF INVENTION: HIV ANTIGEN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolach and Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,510  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/985,949  
FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 216-309P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 500 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
AMTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
US-08-375-510-1  
Alignment Scores:  
Pred. No.: 6,92e-130 Length: 500  
Score: 2249.00 Matches: 421  
Percent Similarity: 92.22% Conservative: 41  
Best Local Similarity: 84.03% Mismatches: 29  
Query Match: 81.13% Indels: 10  
DB: 1 Gaps: 5  
US-09-475-704A-3 (1-1479) x US-08-375-510-1 (1-500)  
QY 1 ATGGGCGCCGCGCCGATCTGCGCGGCGGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
DB 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuSerLysTrpGluLysIleArg 20  
QY 61 CTGGCGCCGCGCGGCAAGAGTGTTCATGATGAAGACACCTGTGTGGCCAGCCGCGAG 120  
DB 21 LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValTrpAlaSerArgIle 40  
QY 121 CTGGAGAAATTTGGCCTGAACCCCGGCTGCTGGAGAACCGAGGGCTGGCAAGCATC 180  
DB 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGlnIleCysArgGlnIle 60  
QY 181 ATCCGCACTGCAACCCGCTGAGCGCGGCGAGCGGAGGCTGAAGAGCTGTTCAC 240  
DB 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnIleLeuArgSerLeuTyrAsn 80  
QY 241 ACCGTGGCAACCTGTATCTGCTGCAAGAAAGATGAGTCCGCGACCAAGAGGCC 300  
DB 81 ThrIleAlaValLeuTyrCysValHisGlnArgIleAspValLysPheThrLysGlnIle 100  
QY 301 CTGGACAAGATCGAGAGAGCAAGAACAGTGCACAGAAATCTCAGAGCGCGAGGCC 360

DB 101 LeuAspLysIleGluGlnIleGlnLysSerLysLysValAlaGlnIleAla--Ala 119  
QY 361 GCCGCAAGAGGC-----AAGGTAGCCGAACTACCCCATGTGCGAAGCTGCGAG 411  
DB 120 AlaAspThrGlyAsnAsnSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGln 139  
QY 412 GCGCAGATGTGCACCGAGCCATCAAGCCCGCGACCCCTGAACGCTGGGTGAAGTATC 471  
DB 140 GlyGlnMetValHisGlnIleAlaIleSerProArgThrLeuAsnAlaTrpValLysVal 159  
QY 472 GAGGAGAGGCTTCAGCGCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGCGCGC 531  
DB 160 GlnGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGlnIleGlyAla 179  
QY 532 ACCCCGCAAGACCTGAACAGATGTTGAACCGTGGCGGCGCACAGCGCCCATGCGAG 591  
DB 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaIleMetGln 199  
QY 592 ATGTGAMGACACATCAACGAGAGCGCGCGAGTGGAGACCGGTGACCCCGTGCAC 651  
DB 200 MetLeuLysGluThrIleAsnGlnIleAlaIleArgIleProArgLeuHisIleProValHis 219  
QY 652 GCGGCGCCCATGGCCCCGCGCGCAAGATGGCGCGCGCGCGCGAGGACATGCGCGCAC 711  
DB 220 AlaGlyProIleAlaProGlyGlnMetArgIleProArgLysSerAspIleAlaGlyThr 239  
QY 712 ACCACACCCCTGACGAGAGATCGCTGATGACAGAACCCCGCTCCCGCTGGGCG 771  
DB 240 ThrSerThrLeuGlnGlnIleGlyTyrMetThrHisAsnProProIleProValGly 259  
QY 772 GACATCTACAGCGGTGATCATCTGAGCTGAAACAAGATGTCGAGATGTACAGCCC 831  
DB 260 GlnIleTyrLysArgTrpIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279  
QY 832 GTGACATCTGAGATCAAGAGAGCGCCCAAGAGCCCTCCCGACTACCTGAGACCG 891  
DB 280 ThrSerIleLeuAspIleArgGlnIleProLysGlnProPheArgAspTyrAlaAspArg 299  
QY 892 TTCCTCAAGACCTCGCGCGCGCGAGAGACACCCGAGAGTGAACATGTGATGACCGAC 951  
DB 300 PheTyrLysThrLeuArgAlaGlnIleAlaSerGlnIleValLysAsnTrpMetThrGlu 319  
QY 952 ACCGTGTGTGTCAGAACGCGCAACCCCGATGTCAAGACATCTGCGGCTTCGAGCCC 1011  
DB 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339  
QY 1012 GCGCGCAGCTGAGAGATATGACCGCTGCGCGGCGCTGGGCGCCCAAGCAAG 1071  
DB 340 GlyAlaThrLeuGlnGlnMetMetThrAlaCysGlnIleValGlyProGlyLysLys 359  
QY 1072 GCCCGCGTGTGCGCGGCGATGAGCGAGCGCAAC-----ACAGGCGTGTGATGCG 1125  
DB 360 AlaArgValLeuAlaGlnIleAlaMetSerGlnIleThrAsnProAlaThrIleMetIleGln 379  
QY 1126 AAGAGCAATTCAGAGCGCCCGCGGCGCATGTCAAGTCTTCACTGCGCGCAAGAGGCG 1185  
DB 380 LysGlyAsnPheAlaArgAsnGlnArgLysThrValLysCysPheAsnCysGlyLysGlnIle 399  
QY 1186 CACATCGCGCGCAATGCGCGCGCGCCCGCAAGAAAGGCTGTGAAAGTGGCGCAAGAG 1245  
DB 400 HisIleAlaLysAsnCysArgAlaProArgLysGlyCysTrpLysCysAspLysGln 419  
QY 1246 GCGCACAGATGAAGAATGACGCGAGCGCGCAACTTCTGGGCAAGATCTGGCGCC 1305  
DB 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439  
QY 1306 AGCCACAAGGCGCGCGCGCAATCTCTGCAAGAGCGCGCCAGCCAGCGCGCGCC 1365  
DB 440 SerHisLysGlyArgProGlyLysAsnPheLeuGlnSerArgProGluProThrAlaProPro 459  
QY 1366 GCGGAGACTTCGCGCTTC-----GAGAGACCAACCCCGCGCAAGACGAGAGCAAG 1419

```

Db      460  Glu1IuSerPheArgPheGlyGluInuThrThrThrProSerGlnIysGlnIuProIle 47
Oy      1420  GACCGCGAG-----ACCTGACCAAGCTGTGGGACAGACCCCTTGAC 14
          |||::|||          |||::|||::|||::|||::|||::|||
Db      480  AspIysGluIuSerProLeuAlaSerLeuArgSerIuPheGlySerAspProSerSer 49
Oy      1474  CAG 1476
          |||
Db      500  Gln 500

RESULT 5
US-08-487-657-1
; Sequence 1, Application US/08487657
; Patent No. 5834267
; GENERAL INFORMATION:
; APPLICANT: Saito, Atsushi
; APPLICANT: Sinagawa, Hideo
; APPLICANT: Nakata, Atsuo
; TITLE OF INVENTION: HIV ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolsch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,657
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,510
; FILING DATE: 18-JAN-1995
; APPLICATION NUMBER: US/07/985,949
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 216-309P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
US-08-487-657-1

Alignment Scores:
Pred. No.:          6.92e-130          Length:          500
Score:              2249.00            Matches:          421
Percent Similarity: 92.22%             Conservative:     41
Best Local Similarity: 84.03%           Mismatches:      29
Query Match:        81.13%             Indels:          10
DB:                  2                  Gaps:             5

US-08-475-704A-3 (1-1479) x US-08-487-657-1 (1-500)
Oy      1  ATGAGCGCGCGCGCAGCATTCTGCGCGCGCGAGAGCTGAGCGCTTGGAGCGCATTCGCG 60

```

Ds	1	MetGlyAlaArgAlaSerValLeuSerGlyValGlyGluMetAspLysTrpGluValIleArg	20
Qy	61	CTGGCGCCCGGCGCAAGAGTGTCTATCATGTATGAACGACCTGGTGGGCGACGGCGAG	120
Ds	21	LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValTrpAlaSerArgLys	40
Qy	121	CTGAGAAAGTTGGCCCTGTAACCCCGGCTGCTGAGAGACCAGCGAGGCTGCACAGCATC	180
Ds	41	LeuGlnTrpPheAlaValAlaAsnProGlyLeuLeuGlnTrpSerGlnGlyCysArgGlnIle	60
Qy	181	ATCGCGCAGCTGCACCCCGGCTCGAGACCGGACCGGAGGCTGAAGACGCTGTCAAC	240
Ds	61	LeuGlnLeuLeuGlnProSerLeuGlnThrAlaSerGlnGluLeuAsnSerLeuLysAsn	80
Qy	241	ACCGTGGCCACCTGTACTGCTGACACGAGATCGAGGTCCGGACACCAAGAGAGCC	300
Ds	81	ThrIleAlaValLeuLysTrpCysValHisGlnArgIleAspValLysAspThrLysGlnAla	100
Qy	301	CTGACAAAGATCGAGAGAGAGACAGAACAAAGTCCAGCAAGAAATCATCAGAGGCGAGCC	360
Ds	101	LeuAspLysIleGlnGlnGlnGlnAlaAsnLysSerLysLysAlaGlnGlnAla--Ala	119
Qy	361	GCCGACAAAGGCG-----AAGTGAAGCCAGAACTACCCCATGTGTCAGAACTCGAC	411
Ds	120	AlaAspThrGlyAsnAsnSerGlnValSerGlnAsnLysTrpProIleValGlnAsnLeuGln	139
Qy	412	GGCAGATGTGTGACACGAGCCATCAGCCCGGACCCCTGAACGCTGGGTGAAGGTGATC	471
Ds	140	GlyIleMetValHisGlnAlaIleSerProArgThrLeuMetAlaTrpValLysValVal	159
Qy	472	GAGAGAAAGGCTTCAGGCTCCGAGGTGATCCCATGTTCAACCGGCTTGAAGGAGGCGCC	531
Ds	160	GlnGlnLysAlaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnGlyAla	179
Qy	532	ACCCCGCAGGACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGGCGGCATGTGAG	591
Ds	180	ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGln	199
Qy	592	ATGCTGAAGACACCATCAACGAGAGAGGCGCCCGATGGGAGACCGGCTACACCCCTGCAC	651
Ds	200	MetLeuLysGlnThrIleAsnGlnGlnAlaIleGlnTrpAspArgLysHisIleProValHis	219
Qy	652	GCCGCGCCCATGCGCCCGGCGCAATGCGAGCGGAGCCCGGCGAGGACATGCGCGGACCC	711
Ds	220	AlaGlyProIleAlaProGlyGlnMetArgLysProAspGlySerAspIleAlaGlyThr	239
Qy	712	ACGAGCACCTTCGAGAGACATGCGCTGTGATGACACCAACCCCCCATCCCCGTGGGC	771
Ds	240	ThrSerThrLeuGlnGlnGlnIleGlyTrpMetThrHisAsnProProIleProValGly	259
Qy	772	GACATCTCAACGCGGTGATCATCTTGGGGCTGAACAGATCTGTCGATGACAGCCCC	831
Ds	260	GlnIleLysLysArgTrpIleIleLeuLysLeuMetLysIleValAlaMetLysSerPro	279
Qy	832	GTCAGCATCTTGACATCAACGAGGCGCCCAAGAGCCCTTCGGGACTACGTGAGCGCC	891
Ds	280	ThrSerIleLeuAspIleArgGlnGlyProLysGlnProPheAspArgPheLysAspArg	299
Qy	892	TTCTTCAAGACCTTCGCGCGGAGCAGACGACCCAGAGGTGAAGAACTGGATGACCGAC	951
Ds	300	PheLysLysThrLeuArgAlaGlnAlaIleSerGlnGlnValLysAsnTrpMetThrGln	319
Qy	952	ACCGTGGCTGTGAGAAAGCCACCCGACCTGCAAGACATCTTCGCGGCTTCGAGCCCC	1011
Ds	320	ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuValPro	339
Qy	1012	GCGCCAGGCTGTGAGAGATGATGACGCTCTCGACGAGGCTGGGCGGCGCCAGCCACAG	1071
Ds	340	GlyAlaThrLeuGlnGlnMetMetThrLysCysGlnGlyValGlyLysProGlyHisLys	359
Qy	1072	GCCGCGGTGCTGGCGGAGCGATGAGCCAGGCCAAC-----ACCAAGGTGATGATGAC	1127

Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGln 379  
QY 1126 AAGAGCACTTCAGAGGCCCCCGGCATGCTCAATGCTTCACTGCGGCAAGAGGGC 1185  
Db 380 LysGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCyseGlyLysGlnGly 399  
QY 1186 CACATGCGCGCAATGCGCGCGCGCGCGCAAGAGGGCTGTGAAAGTGGCGGCAAGAG 1245  
Db 400 HisIleAlaLysAsnCyseAlaProArgLysLysGlyCysThrLysCysGlyLysGln 419  
QY 1246 GGCACACAGATGAGAGTGCACCGAGCGCCAGCCCACTTCTGGGGCAAGATCTGGCCC 1305  
Db 420 GlyHisGlnMetLysAsnProCysThrGlnArgGlnAlaAsnPheAsnLysIleThrPro 439  
QY 1306 AGCCACAAAGGCGCGCGCGCGCACTTCTGCGAGAGCGCGCGCGCAAGCGCGCGCC 1365  
Db 440 SerHisLysGlyArgProGlnAsnPheLysSerArgProGlnProThrAlaProPro 459  
QY 1366 GCCGAGAGCTTCCGCTTC-----GAGAGACCAAGCGCGCGCGCAAGAGAGAGAG 1419  
Db 460 GlnGlnSerPheArgPheGlnGlnGlnThrThrProSerGlnLysGlnGlnProIle 479  
QY 1420 GACCGCGAG-----ACCTGACCAAGCTGAGAGAGCTGTTCGCAAGCAAGCGCGTGAAC 1473  
Db 480 AspLysGlnLeuTyrrProLeuAlaSerLeuArgSerLeuPheGlySerAspProSerSer 499  
QY 1474 CAG 1476  
Db 500 Gln 500  
RESULT 6  
US-09-309-572-16  
/ Sequence 16, Application US/09309572  
/ Patent No. 6440730  
/ GENERAL INFORMATION:  
/ APPLICANT: Heinrich-Pette-Institut  
/ TITLE OF INVENTION: Retroviral hybrid vectors pseudocyped with LCMV  
/ FILE REFERENCE: P50489  
/ CURRENT APPLICATION NUMBER: US/09/309,572  
/ EARLIER FILING DATE: 1999-05-11  
/ EARLIER APPLICATION NUMBER: DE 198 56 463  
/ EARLIER FILING DATE: 1998-11-26  
/ NUMBER OF SEQ ID NOS: 24  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 16  
/ LENGTH: 500  
/ TYPE: PRF  
/ ORGANISM: Human immunodeficiency virus type 1  
/ FEATURE:  
/ OTHER INFORMATION: gag polypotein  
US-09-309-572-16  
Alignment Scores:  
Pred. No.: 6,92e-130 Length: 500  
Score: 2249.00 Matches: 421  
Percent Similarity: 92.22% Conservative: 41  
Best Local Similarity: 84.03% Mismatches: 29  
Query Match: 81.13% Indels: 10  
DB: 4 Gaps: 5  
US-09-475-704a-3 (1-1479) x US-09-309-572-16 (1-500)  
QY 1 AAGGCGCGCGCGCGCGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCGC 60  
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGlnLeuAspLysTrpGlnLysIleArg 20  
QY 61 CCGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGACACCTGTGTGGCGCGCGCGAG 120  
Db 21 LeuArgProGlyGlyLysLysGlnTyrrLysLeuLysHisIleValIleThrAlaSerArgGln 40  
QY 121 CTGGAAGAGTTCGCGCTGGAACCGCGCGCTGTGAGAGAGAGAGAGCGGCTGCAAGCATC 180  
Db 41 LeuGlnArgPheAlaValAsnProGlyLeuLeuGlnThrSerGlnLysCysArgGlnIle 60

QY 181 ATCCCGCAGCTGACACCCCGCTGCAAGCCGCGACGAGAGCTGAAGAGCTGTTCAC 240  
Db 61 LeuGlnGlnLeuGlnProSerLeuGlnThrGlySerGlnGlnLeuAspSerLeuTyrrAsn 80  
QY 241 ACCGTGACACCTGTACTGCTGCGAGAGAAAGATCTGAGGTCCGCGACCAAGAGAGCC 300  
Db 81 ThrIleAlaValLeuTyrrCysValHisGlnArgIleAspValLysAspThrLysGlnAla 100  
QY 301 CTGCAAGAGTGAAGAGAGAGAGCAAGATGTCACAGAAATTCACAGCGCGCGAGCC 360  
Db 101 LeuAspLysIleGlnGlnGlnLysSerLysLysValGlnGlnAla--Ala 119  
QY 361 GCCGCAAGAGGC-----AAGTGAAGCCAGAACTACCCATGTGTGACAGACTGCA 411  
Db 120 AlaAspThrGlyAsnAsnSerGlnValSerGlnAsnTyrrProIleValGlnAsnLeuGln 139  
QY 412 GCGCAGATGATGACACAGGCGCATGAGCCCGCGACCTGAAAGCGCTGGGTGAAGGTATC 471  
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValAl 159  
QY 472 GAGAGAAAGCCTTGAGCGCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGCGCC 531  
Db 160 GlnGlnLysAlaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnLysAla 179  
QY 532 ACCCGCGAGACTGAACAGCATGTTGAACACCGTGGCGCGCACAGCGCGCATGCA 591  
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGln 199  
QY 592 ATGCTGAAGACACCATCAACGAGAGCGCGCGAGTGGAGCGCGTGCACCGCTGCAC 651  
Db 200 MetLeuLysGlnThrThrIleAsnGlnGlnAlaAlaGlnThrPaspArgLeuHisProValHis 219  
QY 652 GCGGCGCGCGCGCGCGCGCGCGAGTGGCGCGAGCGCGCGCGAGAGCATGCGCGGAC 711  
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGlnProArgLysSerAspIleAlaGlyThr 239  
QY 712 ACCAGACCGCTGAGAGAGAGATGCGCTGATGACAGCAAGCGCGCGCATCCCGTGGC 771  
Db 240 ThrSerThrLeuGlnGlnGlnIleGlyTyrrMetThrHisAsnProProIleProValGly 259  
QY 772 GACATCTCAAGCGGTGATCATCTGGGCGCTGAAGATGTCGGATGATGACGCGCC 831  
Db 260 GlnIleTyrrLysArgTrpIleIleLeuGlyLeuAsnLysIleValAlaArgMetLysSerPro 279  
QY 832 GTGACATCTGAGACATCAAGAGAGCGCGCGCGCGCGCGCGCGCTTCCGAGTACGTGACGCG 891  
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGlnProPheArgAspTyrrValAspArg 299  
QY 892 TTCTTCAAGACCTGCGCGCGCGAGAGAGACCGCGAGGTGAAGACTGATGACCGAC 951  
Db 300 PheTyrrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnTrpMetThrGln 319  
QY 952 ACCGTGATGATGACAAAGCGCAAGCGCGAGTGAAGACATCTCGCGGCTGTGGCGCC 1011  
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuIlePro 339  
QY 1012 GCGCGCGCGCGCGAGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAAG 1071  
Db 340 GlyAlaThrLeuGlnGlnGlnMetCethrAlaCysGlnGlnValGlyGlyProGlyHisLys 359  
QY 1072 GCGCGCGCTGCGCGCGAGCGCATGAGCGCAAGCGCGCAAG-----ACCAAGGTGATGATGAG 1125  
Db 360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnProAlaThrIleMetIleGln 379  
QY 1126 AAGAGCACTTCAGAGGCCCCCGGCATGCTCAATGCTTCACTGCGGCAAGAGGGC 1185  
Db 380 LysGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCyseGlyLysGlnGly 399  
QY 1186 CACATGCGCGCAATGCGCGCGCGCGCACTTCTGCGAGAGCGCGCGCGCAAGCGCGCGCC 1245  
Db 400 HisIleAlaLysAsnCyseAlaProArgLysLysGlyCysThrLysCysGlyLysGln 419



```

Db      440 SerHisIySgIyArProGluYAsnPhnLeuGlnSerArgProGluProThrAlaProPro 459
QY      1366 GCCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGGCGAGAGAGAGAGAG 1419
Db      460 GlnGluSerPheArgPheGlyGlnGluThrThrProSerGlnYsgGlnGluProIle 479
QY      1420 GACCGGAG-----ACCTGACACCGCTGGAAGAGCTGTCGCAACGACCCCTGAGC 1473
Db      480 AspYsGlnLeuYrProLeuAlaSerLeuArgSerLeuPheGlySerAspProSerSer 499
QY      1474 CAG 1476
Db      500 Gln 500

RESULT 8
US-09-952-060-35
; Sequence 35, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emni, Emilio A.
; APPLICANT: Youli, Kima
; APPLICANT: Betz, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV-1-GAG, POL, NEF AND
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimized gag-1A pol fusion
US-09-952-060-35

Alignment Scores:
Pred. No.: 1,44e-129 Length: 1350
Score: 2245.00 Matches: 419
Percent Similarity: 92.00% Conservative: 41
Best Local Similarity: 83.80% Mismatches: 32
Query Match: 80.99% Indels: 8
DB: 4 Gaps: 4

US-09-475-704a-3 (1-1479) x US-09-952-060-35 (1-1350)
QY      1 ATGGGCGCCCGCGCCGACGATCTGCGCGCGGCGGACGCTGAGCGCTGGAGCGGATCCGC 60
Db      1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspLysTrpGluYsIleArg 20
QY      61 CTGGCGCCCGCGGCGGCAAGAGTGTACATGATGAAGACACCTGGTGTGGCGCGAGCGCGAG 120
Db      21 LeuArgProGlyGlyYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYs 40
QY      121 CTGGAAGATTCGCGCTGGAACCCCGGCTGCTGGAGACGAGGAGGCTGCAAGAGATC 180
Db      41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGlnGlyCysArgGlnIle 60

```

```

QY      181 ATCCCGCAGCTGCACCCCGCTCGACAGCCGCGAGAGAGCTGAAGAGCTGTTCAAC 240
Db      61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuYrAsn 80
QY      241 ACCCGGCGCACCCTGACTGCGTGCACGAGAGATCGAGGCTCGCGAGACCAAGAGGCG 300
Db      81 ThrValAlaThrLeuYrCysValHisGlnYsYsIleAspValLysAspThrYsGluAla 100
QY      301 CTGACACAGATCGAGAGAGAGAGCAAGATGTCACAGAGATCCAGAGGCGCGAGGCC 360
Db      101 LeuGluYsIleGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
QY      361 -----GCCGACAAAGGCAAGTGAAGCCAGAACTACCCCATGCTGACAGAACTGCAAGGCC 414
Db      121 GlyThrGlyAsnSerSerGlnValSerGlnAsnYrProIleValGlnAsnLeuGlnGly 140
QY      415 CAGATGTGTCACAGAGCCATCAGCCCGCGACCCCTGGAAGCGCTGGAGTGAAGTGAACG 474
Db      141 GlnMetValHisGlnAlaIleSerProMgThrLeuAsnAlaTrpValLysValValGlu 160
QY      475 GAGAAAGCCTTGAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGCGCGCAC 534
Db      161 GlnYsAlaPheSerProGluValIleProMetPheSerAlaLeuSerGlnGlyAlaThr 180
QY      535 CCCGAGAGCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCGCGCATGCAATG 594
Db      181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMet 200
QY      595 CTGAAGAGACCATGAAGAGAGAGCGCGCGAGTGGAGCCCGTGCACCCGTCAGACCC 654
Db      201 LeuYsGluThrIleAsnGlnGlnAlaGlnAlaGlnTrpAspArgLeuIleProValHisAla 220
QY      655 GCGCCCATCGCCCGCGCAGATGCGCGAGCGCCCGCGAGAGATCGCGCGAGCGCGCAC 714
Db      221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY      715 AGCACCTGCGAGAGACGATCGCTGATGACAGCAACCCCGCATCCCGTGGCGGAC 774
Db      241 SerThrLeuGlnGlnGlnIleGlyTrpMetThrAsnAsnProProIleProValGlyGlu 260
QY      775 ATCTCAAGCGGTGATCATCTGCGCGCTGAACAAGATCGGCGGATGATACAGCCCGCG 834
Db      261 IleYrYsArgGlyTrpIleLeuGlnYleuAsnYsIleValArgMetYsSerProThr 280
QY      835 AGCATCTGACATCAAGAGAGGCGCCCAAGAGCGCTTCGCGCATAGTGAACCGCTTC 894
Db      281 SerIleLeuAspIleArgGlnGlyProYsGluProPheArgAspYrValAspArgPhe 300
QY      895 TTCAAGACCTGCGCGCGCGAGAGACGACCCAGAGAGGTGAAGACTGATGACCGACACC 954
Db      301 TyrYsThrLeuArgAlaGlnGlnAlaSerGlnGlnValYsAsnTrpMetThrGluThr 320
QY      955 CTGCTGTGGCAAGAGCGCAACCCCGCATGCAAGACATCTCGCGGCTCGCGCCCGCG 1014
Db      321 LeuLeuValGlnAsnAlaAsnProAspYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYs 340
QY      1015 GCCAGCTGAGAGATGATGACCGCTGCGAGGCGGTGGCGCGCCGACGACCAAGAGGCC 1074
Db      341 AlaThrLeuGlnGlnGlnMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisYsAla 360
QY      1075 CGCGTGTGGCGAGGCGCATAGCGCAACACACAC-----GTATGATGACAGAG 1128
Db      361 ArgValIleuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380
QY      1129 AGCACTTCAAGAGGCGCGCGCGCATCTGCAAGTCTTCAACTGCGCGAGAGAGGCGCAC 1188
Db      381 GlyAsnPheArgAsnGlnArgYsThrValYsYsYsPheAsnYsGlyLysValGlyHis 400
QY      1189 ATGCGCCGCAACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCGCAAGAGGCGC 1248
Db      401 IleAlaYsAsnYsArgAlaProArgYsYsGlyCysTrpYsYsCysGlyLysGlnGly 420
QY      1249 CACCAAGTGAAGAGACTGACCGAGCGCGCAAGCGCACTTCTGGGCGAGATCTGCGCGCAC 1308

```

```

|||||
Db      421 HtAGlMetLysAspGysaengluatrglnAlaAnPheLenglyLysrletrProSer 440
      1309 CACAAGGGGGCCCCGGCAACTTCCTGCAGAGACGGCCCGGACCCACCCCGCCGCC 1368
      441 HtLysGlyArProdiLysAnPheLengInSerArxProGluProThrLarProGlu 460
      1369 GAGAGCTTCGGCTT-----GAGAGACACCCCGCGCCAGAGAGAGAGAGAGAC 1422
      461 GluSerPheArPheGlygluGluLysrtrrtrProSerGlnLysGlnLurProleap 480
      1423 CGCGAG-----ACCCTGACCAAGCTGGAAGAGCTGTGGAGACGACCCCTGAGCCAG 1476
      481 LysGluLeuTyLProLeuAlaSerLeuArxSerLeuPheGlyLysAnspProSerSerGln 500

```

```

RESULT 9
US-08-463-210-8
/ Sequence 8, Application US/08463210
/ Patent No. 6001977
/ GENERAL INFORMATION:
/ APPLICANT: CHANG, Nancy T.
/ APPLICANT: GALLO, Robert C.
/ APPLICANT: WONG-STALD, Floesie
/ TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morgan & Finnegan, L.L.P.
/ STREET: 345 Park Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10154-0053
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Releasee #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/463,210
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 06/693,866
/ FILING DATE: 23-JAN-1985
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 06/659,339
/ FILING DATE: 10-OCT-1984
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Serunian, Leslie A.
/ REGISTRATION NUMBER: 35,353
/ REFERENCE/DOCKET NUMBER: 2026-4193US2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 512 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: HTLV-III
/ NAME/KEY: Protein
/ LOCATION: 1..512
/ OTHER INFORMATION: /note= "gag protein of HTLV-III"
US-08-463-210-8
Alignment Scores:
Pred. No.: 4.99e-129 Length: 512
Score: 2235.00 Matches: 424

```

```

Percent Similarity: 90.64%
Best Local Similarity: 82.65%
Query Match: 80.63%
DB: 3 Gaps: 7
US-09-475-704a-3 (1-1479) x US-08-463-210-8 (1-512)
      1 ATGGGCGCCCGGCGAGCATCTCGCGCGGCAAGTGCAGCGCTGGGAGCGCATCCGC 60
      1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuMaBPAgTgPduLysrleArg 20
      21 LeuArxProGlyGlyLysLysLysLysrlyrLysLeuLysrleValAlaTrpLaserArgLyu 40
      61 CTGCGCCCGCGCGGCAAGTGTCTACATGTAAGCACTGTGTGTGGCCAGCGCGAG 120
      21 LeuArxProGlyGlyLysLysLysLysrlyrLysLeuLysrleValAlaTrpLaserArgLyu 40
      121 CTGGAAGAAGTTCGCGCTTAACCCCGGCTGTGCAGACACCGAGGCTGCAGACGATC 180
      41 LeuGluArxPheAlaValAlaAnPProdiLysLeuGluLurtrrSerGluGlyCyArxGlnIle 60
      181 ATCCGCAAGCTGCACCCCGCCCTGCAGACCGGCGAGAGAGAGCTGAAGAGCTGTCAAC 240
      61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnLurLueArxSerLeuTyLrAn 80
      241 ACCGTGCCACCTGTACTGTGCTGCAGAGAAGATCGAGTTCGCGACACCAAGAGGCC 300
      81 ThrValAlaTrnLeuTyLysValAlaGlnArxGlnIleGlyLysrtrLysrleGluAla 100
      301 CTGGAAGAAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
      101 LeuArxPheLysrleGluGluGluGlnLysrlyrLysLysrlyrLysrlyrLysrlyr 119
      361 GCCGCAAGAGGC-----AAGGTGACCAAGAACTACCCCATCTGTGCAGAACTCGAG 411
      120 AlaAspTrnGlyHisSerSerGlnValSerGlnAnrTyLProIleValGlnAnrIleGln 139
      412 GCGCATGTGTGCACCAAGGCAATCAGCCCGGCACTGGAAGCTGTGGTGAAGTATC 471
      140 GlyLysMetValHisGlnAlaIleSerProArxTrLysLeuAnAlaTrpValLysVal 159
      472 GAGGAAGAGGCTTCAGCGCCGAGGTGATCCCATGTTCACCGGCTGAGGAGGCGCC 531
      160 GlnGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGlnIlyAla 179
      532 ACCCCCGAGCACTGTAACAGCATGTGTAACACCGTGGCGGCGCACAGCGCCATGCGAG 591
      180 ThrProGlnMaBPAleuAnrThrMetLeuAnrThrValGlyGlnHisGlnAlaIleMetCln 199
      592 ATGTGAAGACACCATTAACGAGAGAGCGCGCGAGTGGAAACCGGTGCACCCGCTGCAC 651
      200 MetLeuLysGluThrIleAsnGluGluAlaIleGluTrpArxArgValHisProValHis 219
      652 GCGGCGCCCATCGCGCCCGGCGAGATGCGCGAGCGCGGCGAGGAGATCGCGGCGAGC 711
      220 AlaGlyProIleAlaProGlyGlnMetArxGluProArxGlySerArPheAlaGlyThr 239
      712 ACCAGACACCTGCAGAGACAGATCGCTGTGATGACAGCAACCCCGCATCCCGTGGGC 771
      240 ThrSerTrnLeuGlnGlnGlnIleGlyTrpMetThrAnAnPProIleProValGly 259
      772 GACATCTACAGCGGTGATCATCTGGGCTGTGAACAAAGATCGTGGATGTACAGCCC 831
      260 GlnIleTyLysArxTrpIleIleLeuGlyLeuAnrLysrleValArxMetLysrPro 279
      832 GTGAGCATCTGCAGACATCAAGAGAGCGCGCGCAAGAGCGCTTCGGGATCTAGTGAAGCG 891
      280 ThrSerIleLeuArPheArxGlnIlyProLysGluProPheArxArgTyLValAspArg 299
      892 TTCTTCAAGACCTTCGCGCGCGAGAGAGAGACCCGAGAGTGAAGAGTGAAGTGAAGCAGAC 951
      300 PheTyLysrtrLysrtrLysrtrLysrtrLysrtrLysrtrLysrtrLysrtrLysrtr 319
      952 ACCGTGCTGTGCAGAGAGCGCAACCCCGATGCAAGACCATCTGCGGCTGTGCGGCC 1011

```



```

Db      320 ThrLeuLeuValGlnAsnAlaenProAspCysLysThrIleuLysAlaLeuGlyPro 339
QY      1012 GGGCCAGCCTGAGAGATGATGATACCGCTGCGAGGGGTGGCGGCCCGCCAGCAG 1071
        |||:::|||||
Db      340 AlaAlaThrLeuGlnGlnMetMetThrAlaCysGlnGlyAlaGlyProGlyHisLys 359
QY      1072 GCGCGGCTCTGGCCGAGCGATGATGACAG---GCCACAC---AGCGATGATGACAG 1125
        |||:::|||||
Db      360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
QY      1126 AAGAGCACTTCAAGAGCGCCCGCGCATGTCAGTGTCAACTGCGGCAAGAGAGGC 1185
        |||:::|||||
Db      380 ArgGlyAsnPheArgSerGlnArgLysMetValLysCysPheAsnCysGlyLysGln 399
QY      1186 CACATGCGCGGCACTGCGCGCGCGCGCGCGAGAGAGGGTGTGGAAGTGGCGAGAG 1245
        |||:::|||||
Db      400 HisThrAlaArgSerCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGln 419
QY      1246 GGCACACAGTGAAGAGTGCACCGAGCGCGAGCGCAACTCTGCGGCAAGATCTGGCC 1305
        |||:::|||||
Db      420 GlyHisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439
QY      1306 AGCCACAGAGCGCGCGCGCGCAAC-----1329
        |||:::|||||
Db      440 SerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGlnProThrAlaProPro 459
QY      1330 TTCCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1383
        |||:::|||||
Db      460 PheLeuGlnSerArgProGlnProThrAlaProGlnGlnSerPheArgSerGlyVal 479
QY      1384 GAGAGAGACCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
        |||:::|||||
Db      480 GlnThrThrThrProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 499
QY      1438 AGCTGAGAGCGCTGTTCGCAACGACCGCTGAGCGCAG 1476
        |||:::|||||
Db      500 SerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512
        |||:::|||||

RESULT 10
US-09-124-900-2
/ Sequence 2, Application US/09124900
/ Patent No. 6268484
/ GENERAL INFORMATION:
/ APPLICANT: KATINGER, Hermann
/ APPLICANT: BUCHACHER, Andrea
/ APPLICANT: ERNST, Wolfgang
/ APPLICANT: BALTAUN, Claudia
/ APPLICANT: FURTSCHER, Martina
/ APPLICANT: TRKOJA, Alexandra
/ APPLICANT: PREDL, Renate
/ APPLICANT: SCHMATTZ, Christel
/ APPLICANT: KLIMA, Annelies
/ APPLICANT: MUSTER, Thomas
/ APPLICANT: STEINDL, Franz
/ TITLE OF INVENTION: HIV-Vaccines
/ FILE REFERENCE: 1939-112B
/ CURRENT APPLICATION NUMBER: US/09/124, 900
/ CURRENT FILING DATE: 1998-07-30
/ PRIOR APPLICATION NUMBER: PCT/EP95/01481
/ PRIOR FILING DATE: 1995-04-19
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 2
/ LENGTH: 512
/ TYPE: PRT
/ ORGANISM: Human immunodeficiency virus type 1
US-09-124-900-2

Alignment Scores:
Pred. No.: 4.99e-129 Length: 512
Score: 2235.00 Matches: 424
Percent Similarity: 90.64% Conservative: 41
Best Local Similarity: 82.65% Mismatches: 26

```

```

Query Match: 80.63% Indels: 22
DB: 3 Gaps: 7
US-09-475-704a-3 (1-1479) x US-09-124-900-2 (1-512)

QY      1 ATGGCGCGCGCGCGAGCATCTGCGCGCGCGCGAGAGCGCTGGAGAGCGATCCGC 60
        |||:::|||||
Db      1 MetGlyAlaArgAlaSerValLeuSerGlyGlyIleuLeuAspArgTrpGlnLysIleArg 20
QY      61 CTGCGCGCGCGCGCGAGAGAGTGTACTATGATGAAGACCTGGTGGCGCGCGAG 120
        |||:::|||||
Db      21 LeuArgProGlyGlyLysLysLysThrLysLeuLysHisIleValTrpAlaSerArgGln 40
QY      121 CTGAGAGAGTGGCGCTGAACCCCGCGCTGCTGAGACACGAGCGAGCGCTGCAAGCATC 180
        |||:::|||||
Db      41 LeuGlnArgPheAlaValAlaAsnProLysLeuLeuGlnTrpSerGlnGlyCysArgGlnIle 60
QY      181 ATCCGCGAGCTGCACCGCGCGCTGACAGCCGCGAGCGAGAGTGAAGAGCGCTGTCAAC 240
        |||:::|||||
Db      61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGlnLeuArgSerLeuTyrAsn 80
QY      241 ACCGTGCGCACCGCTGACTGCGTGCACGAGAGATGAGTCCGCGACACCAAGAGAGCC 300
        |||:::|||||
Db      81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGlnIleLysAspThrLysGlnAla 100
QY      301 CTGACCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
        |||:::|||||
Db      101 LeuAspLysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 119
QY      361 GCCGACAGAGGCG-----AAGGTGAGCGCAAGACTCCCATGCTGCAGAACCTGAG 411
        |||:::|||||
Db      120 AlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProLysAlaGlnAsnIleGln 139
QY      412 GGCAGATGTGTCACAGAGAGCATGACCCCGCGACCTGAGAGCGCTGGGTGAAGTGTATC 471
        |||:::|||||
Db      140 GlyIleMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysVal 159
QY      472 GAGAGAGAGCGCTTACCGCGAGAGTATCCCATGTTCCACCGCGCTGAGAGAGAGCGCC 531
        |||:::|||||
Db      160 GlnGlnLysAlaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnIleVal 179
QY      532 ACCCGCGAGAGCTGAACAGATGTGAACCGCGCGCGCGAGCGAGCGCGCGAGCGAG 591
        |||:::|||||
Db      180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGln 199
QY      592 ATGCTGAAGAGACACCATCAAGAGAGAGCGCGCGAGTGGAGACCGCGTGCACCGCTGAC 651
        |||:::|||||
Db      200 MetLeuLysGlnThrIleAsnGlnGlnAlaIleGlnTrpAspArgValHisProValHis 219
QY      652 GCGGCGCGCATCGCGCGCGCGAGTGGCGAGCGCGCGCGCGAGCGAGCATGCGCGAGCC 711
        |||:::|||||
Db      220 AlaGlyProLysAlaProGlyGlnMetArgGlnProArgLysSerAspIleAlaGlyThr 239
QY      712 ACCAGCAGCGCGAGAGAGAGATGCGCGTGAAGACACCAACCGCGCGCGCGCGTGGGC 771
        |||:::|||||
Db      240 ThrSerThrLeuGlnGlnGlnIleLysTyrMetThrAsnAsnProProIleProValGly 259
QY      772 GACATCTCAAGCGCGTGTATCTCTGGCGCTGGAACAAGATGTCGCGATGACAGCGCC 831
        |||:::|||||
Db      260 GlnIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerPro 279
QY      832 GTAGCATCTCGAGCATCAAGAGAGCGCGCGAGAGAGCGCTTCCGAGTCTAGTGAAGCGC 891
        |||:::|||||
Db      280 ThrSerIleLeuAspIleArgGlnLysProLysGlnProPheArgAspTyrValAspArg 299
QY      892 TTCTTCAAGACCGCGCGCGCGAGAGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 951
        |||:::|||||
Db      300 PheTyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnTrpMetThrGln 319
QY      952 ACCCTGCTGTGTCAGAGAGCGCAACCGCGAGTGCAGAGACATCTGCGCGCTCTGCGCGCC 1011
        |||:::|||||
Db      320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339

```

```

QY 1012 GCGCCAGCCTGAGAGATGATGACCGCTGCGAGCGGCCGACCAAG 1071
   |||::|
Db 340 AAlaAAThLeuGlUmetMetThrAlaCySgInGlyValGlyProGlyHisLys 359
   |||::|
QY 1072 GCGCCGCTCTGGCCGAGCGATGAGCCAG--GCCAACAC--AGCTGATGATGAG 1125
   |||::|
Db 360 AAlaGValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrLeuMetGln 379
   |||::|
QY 1126 AAGAGCAATTCAGAGGCGCCGCGCATATCGTCAATGCTTCAATCGCGCAAGAGGGC 1185
   |||::|
Db 380 ATGGAlaAspPheArgAsnGlnIleArgLysMetValLysCysPheAsnGlyLysGly 399
   |||::|
QY 1186 CACATGCGCCGCAACTGCGCGCGCGCCGCAAGAGGCTGTGAAGTGGCGCAAGAG 1245
   |||::|
Db 400 HLeuThrAlaArgAsnGlyArgAlaProArgLysGlyCysThrLysGlyGly 419
   |||::|
QY 1246 GCGCACCATGAGATGACATGACCGGCGCGCGCAACTTCTGGGCAAGATCTGGCC 1305
   |||::|
Db 420 GlyHLeuMetLysAspCysThrGluArgGlnAlaAsnPheLysGlyLysIleThrPro 439
   |||::|
QY 1306 AGCCCAAGGCGCGCGCGCGCAAC----- 1329
   |||::|
Db 440 SerTyrLysGlyArgProGlyAsnPheLysGlnSerArgProGluProThrAlaProPro 459
   |||::|
QY 1330 TTCTGTGACAGCGCGCGCGCGCAACCGCGCGCGCGCGAGCTTCCGCG-----TTC 1383
   |||::|
Db 460 PheLeuGlnSerArgProGluProThrAlaProProGluLysSerPheArgSerGlyVal 479
   |||::|
QY 1384 GAGAGACCAACCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
   |||::|
Db 480 GlnThrThrThrProGlnGlyGlnGluProIleAspLysGlnLeuTyrProLeuThr 499
   |||::|
QY 1438 AGCGTGAAGAGCTGTTCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
   |||::|
Db 500 SerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512
   |||::|

RESULT 11
US-08-463-028-8
; Sequence 8, Application US/08463028
; Patent No. 6610476
;
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: MONG-STAL, Floesie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154-0053
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,028
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/693,866
; FILING DATE: 23-JAN-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/659,339
; FILING DATE: 10-OCT-1984
; ATTORNEY/AGENT INFORMATION:
; NAME: Serutian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4193053
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: HTLV-III
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..512
; OTHER INFORMATION: /note="gag protein of HTLV-III"

US-08-463-028-8

Alignment Scores:
Pred. No.: 4,996-129 Length: 512
Score: 2235.00 Matches: 424
Percent Similarity: 90.64% Conservative: 41
Best Local Similarity: 82.65% Mismatches: 26
Query Match: 80.63% Indels: 22
DB: 4 Gaps: 7

US-09-475-704a-3 (1-1479) x US-08-463-028-8 (1-512)

QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGACGCGCTGGAGCGCATTCGC 60
   |||::|
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspArgTrpGlyLysLeuArg 20
   |||::|
QY 61 CTGGCGCGCGCGCGGAAAGTCTACTATGAAAGCACTGTGTGGCCAGCGCGAG 120
   |||::|
Db 21 LeuArgProGlyGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 40
   |||::|
QY 121 CTGAGAAATTCGCGCTGAACCGCGCGCTGTGAGACGAGAGAGAGAGAGAGAGAGAGAG 180
   |||::|
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGlnThrSerGlnGlyCysArgGlnIle 60
   |||::|
QY 181 ATCCGCGAGCTGACCGCGCGCGCTGACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 240
   |||::|
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGlnLeuArgSerLeuTyrAsn 80
   |||::|
QY 241 ACCGCGCGCGCGCTGACTGCGTGCACGAGAAATCGAGTCCGCGAGACCAAGAGAGCC 300
   |||::|
Db 81 ThrValAlaThrLeuTyrCysValHisGlnIleGlyLysPheThrLysGlnAla 100
   |||::|
QY 301 CTGACCAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
   |||::|
Db 101 LeuAspLysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 119
   |||::|
QY 361 GCGGCAAGGCG-----AAGGTGAGCCAGAACTACCCCATCTGTGACAGACTGCGAG 411
   |||::|
Db 120 AAlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
   |||::|
QY 412 GCGCGAGATGTGACACGAGCATGACCGCGCGCGAGACCGCTGGGTGAAGTATC 471
   |||::|
Db 140 GlyIleMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValAl 159
   |||::|
QY 472 GAGGAGAGGCGCTTCAAGCGCGAGGTATCCCATGTTCACCGCGCTGAGAGAGAGCGCC 531
   |||::|
Db 160 GlnGluLysValAlaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnGlyAla 179
   |||::|
QY 532 ACCCGCGAGACCTGAACAGACATGTTAAACCGTGGCGGCGCACAGCGCGCATGCGAG 591
   |||::|
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGln 199
   |||::|
QY 592 ATGCTGAGACACCATTAAGAGAGAGCGCGCGAGTGGAGCGCGGTGACCGCTGAC 651
   |||::|
Db 200 MetLeuLysGlnThrIleAsnGlnGlnAlaIleGlnTrpAspArgValHisIleProValHis 219
   |||::|
QY 652 GCGGCGCGCGCGCGCGCGAGATGCGCGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
   |||::|

```

```
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerIleAlaGlyThr 239
Qy 712 ACCAGCACTCGAGAGAGAGATGCGCTGATGACCAACCCCGCATCCCGTGGAG 771
Db 240 ThrSerThrLeuGlnGlnIleGlyTyrMetThrAsnAsnProIleProValGly 259
Qy 772 GACATTTACAAGGGGTGATCATCTCGGGCTTAACAAGATCTGGGATGTACAGCCC 831
Db 260 GlnIleTyrLeuArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279
Qy 832 GTGAGATCTGTGACATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 891
Db 280 ThrSerIleLeuAsnIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
Qy 892 TTCTTCAAGACCTGTGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951
Db 300 PheTyrIleThrLeuArgAlaGlnGlnAlaSerGlnGluValLysAsnTyrMetThrGlu 319
Qy 952 ACCCTGTGTGTGAGAACCGGCAACCCGACTGCAAGACCATCTCGCGCTCTCGGCCC 1011
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
Qy 1012 GGGCGAGCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1071
Db 340 AlaAlaThrLeuGlnGlnIleMetMetThrAlaCysGlnGlyValGlyLysProGlyHisLys 359
Qy 1072 GCCCGGCTGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1125
Db 360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
Qy 1126 AAGAGCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1185
Db 380 ArgGlyAsnPheArgAsnGlnIleArgLysMetValLysCysPheAsnCysGlyLysGlnGly 399
Qy 1186 CACATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1245
Db 400 HisThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGln 419
Qy 1246 GGCACACAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1305
Db 420 GlyHisGlnMetLysAspCysThrGlnIleArgGlnAlaAsnPheLeuLysIleTyrPro 439
Qy 1306 AGCCACAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 1329
Db 440 SerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
Qy 1330 TTCCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGCTTCCG 1383
Db 460 PheLeuGlnSerArgProGluProThrAlaProProGluGlnSerPheArgSerGlyVal 479
Qy 1384 GAGAGAGACCAACCCCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 480 GlnThrThrThrProProGlnLysGlnIleProIleAspLysGlnLeuTyrProLeuThr 499
Qy 1438 AGCTTGAAGAGCTGTTCGCAACGAGCCCGCTGAGCCAG 1476
Db 500 SerLeuArgSerLeuPheGlnAsnAspProSerSerGln 512

RESULT 12
US-09-952-060-33
; Sequence 33, Application US/09952060
; Patent No. 673393
; GENERAL INFORMATION:
; APPLICANT: Emili, Emilio A.
; APPLICANT: Youli, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
```

```
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 493
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cpa-gag fusion open reading frame
US-09-952-060-33

Alignment Scores:
Pred. No.: 3,79e-121 Length: 493
Score: 2106.00 Matches: 394
Percent Similarity: 92.29% Conservative: 37
Best Local Similarity: 84.37% Mismatches: 28
Query Match: 75.97% Indels: 8
DB: Gaps: 4

US-09-475-704A-3 (1-1479) x US-09-952-060-33 (1-493)
Qy 100 CTGTGTGGGCGAGCGGAGGCTGTGAGAAAGTTCGCTGAACCCCGGCTGTGAGAGC 159
Db 27 IleValITPAlaSerArgGluLeuGlnArgPheAlaValAsnProGlyLeuLeuIthr 46
Qy 160 ACCGAGGCTGTGAGCAAGATCATCTCGCAAGTGTGACCCCGGCTGTGAGAGCGGAG 219
Db 47 SerGlnGlyCysArgGlnIleLeuGlnIleLeuGlnProSerLeuGlnIthrGlySerGln 66
Qy 220 GAGCTGAGAGGCTTTTCAACACCGGTGGCCACCCCTGTCTGGTGCAGAGAAAGTTCAG 279
Db 67 GlnLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnLysIleAsp 86
Qy 280 GTCCCGCACCAAGAGAGGCGCTGTGACAAAGATCGAGAGAGAGAGAGAGAGAGAGAG 339
Db 87 ValLysAspThrLysGlnAlaLeuGlnLysIleGlnGlnGlnGlnAsnLysSerLysLys 106
Qy 340 AAGATCCAGCAGGCGGAGGCGC-----GCCGCAAGGCGCAAGTGAGCCAGAACTAACCC 393
Db 107 LysAlaGlnGlnAlaAlaAlaGlyThrGlnLysAsnSerSerGlnValSerGlnAsnTyrPro 126
Qy 394 ATCTGTGAGAACTCTGACAGGCGCAAGTGTGTGACACGAGGCGATCAAGCCCGGACCTGAA 453
Db 127 IleValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsn 146
Qy 454 GCGTGGTGAAGGTGTGAGAGAGAGGCGCTTACAGCCCGGAGGTGATCCCATGTTTCAAC 513
Db 147 AlaTyrValLysValValGlnGlnLysAlaPheSerProGluValAlaIleProMetPheSer 166
Qy 514 GCGCTGAGAGAGGCGGCAACCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
Db 167 AlaLeuSerGlnGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGly 186
Qy 574 CACCAAGGCGCGCAGATGCTGAAGACACCAATCAACGAGAGAGAGAGAGAGAGAGAGAG 633
Db 187 HisGlnAlaAlaMetGlnMetLeuLysGlnIleAsnGlnGlnAlaAlaGlnTyrAsp 206
Qy 634 CGCGTGCACCCCGTGTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 693
Db 207 ArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGly 226
```

OY	694	AGCGCATCGCGGGACCCAGCAACCCCTGGACGAGACATGCGCGTGAATGACCGACAC	755
Db	227	SerapillleallaglyThrThrSerThrLeuGlnGlnIleIleLyTTrpMetThrAsn	246
OY	754	CCCCCATCCCGTGGGCGACATCTTACAACGCGGTGATCATCTGGGCGCTGAACAAGATC	813
Db	247	ProProlleProValGlyIuleTyrLybATGTrpIleIleuendlyeunbnySile	266
OY	814	GTGCGGATGTACAGCCCGGTGAGCATCTTGACATAGACAGGGGCCCAAGAGCCCTTC	873
Db	267	ValAGMetTyrSerProThrSerIleLeuAspIleArgGlnGlyProlysgInuProPhe	286
OY	874	CGCGACTACGTGACCGGCTCTTACAAGACCCCTGGCGCGGAGAGAGCAACCCAGAGGNG	933
Db	287	ArgAspTyrValAspArgPheTyrLybTrnLeuArgAlaGlnGlnIleAspGlnGlnuVal	306
OY	934	AAGAACTGGATGACCGGACACCCCTGCTGCTGACAGAACGCCAACCCCGCATCTGCAAGACATC	993
Db	307	LybAsnTrpMetThrGlyThrIleuLeuValGlnIleAsnIleAsnProAspGlybThrIle	326
OY	994	CTGCGCGCTCTGCGCCCGCGCGCCAGCCTGAGAGAGATGATGACGCGCTGCCAGGGCGTG	1055
Db	327	LeuIysAlaLeuGlyProAlaIleAlaThrLeuGlnGlnIleuMetThrAlaCysGlnGlyVal	346
OY	1054	GGCGGCGCCGACGACCAAGGCGCGCGCTGGCGCGGACGAGATGACGACGCGCAACACCAAC	1111
Db	347	GlyGlyProGlyIhIbLysAlaArgValLeuAlaGlnAlaIleMetSerGlnValThrAsnSer	366
OY	1114	-----GTGATGATGCAGAGAAGCAACTTCAAGGGCGCCCGCGCATCGTCAAGTGTCTTC	1165
Db	367	AlaThrIleMetMetGlnArgGlyAsnPheArgAsnGlnArgGlyThrValIysCysPhe	386
OY	1168	AACTGCGGCAAGAGAGGGCCACATCGCCCGCAACTGCGCGCGCCCGCGCAAGAGGGCTGC	1222
Db	387	AsnCysGlyLysValGlyIhIbIleAlaIysAsnCysArgAlaProArgLysIysGlyCys	406
OY	1228	TGGAGTGTCCGCAAGAGGGGCCACAGATGAAGAGATGTGACCGAGGCGCAAGGCAACTTC	1283
Db	407	TrpLysCysGlyLysGlnGlnIhIbGlnMetLysAspCysAsnGlnuArgGlnIleAsnPhe	426
OY	1288	CTGGGCAAGATGTGGCGCCGACCAAGAGGGCGCGCGCGCAACTCTGTGAGAGCGCGCCC	1343
Db	427	LeuGlyLysIleTrpProSerIhIbLysGlyArgProGlyIhAsnPheLeuGlnIleArgPro	446
OY	1348	GAGCCCAACGCGCCCGCGCGCGAGAGCTTCGCGCTTC-----GAGGAGACCAACCCCGGC	1403
Db	447	GluProThrAlaProProGlnGlnuSerPheArgPheGlyGlnGlnuLysThrThrProSer	466
OY	1402	CAGAAAGCAGAGAGCAAGACCGCGCAG-----ACCCTGACACAGCTGAAGAGCTGTTC	1453
Db	467	GlnLysGlnGlnuProIleAspLysGlnuLeuTyrProLeuAlaSerIleuArgSerLeuPhe	486
OY	1456	GGCAACGACCCCTGAGCCAG 1476	
Db	487	GlyAsnAspProSerSerGln 493	
RESULT 13			
US-08-127-499A-11			
Sequence 11, Application US/08127499A			
Patent No. 5510264			
GENERAL INFORMATION:			
APPLICANT: VAN ALSTYNE, Diane			
APPLICANT: SHARMA, Lawrence Rajendra			
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED			
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES			
NUMBER OF SEQUENCES: 40			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Foley & Lardner			
STREET: 3000 K Street, N.W., Suite 500			
CITY: Washington			
STATE: D.C.			
COUNTRY: USA			
ZIP: 20007-5109			

```

COMPUTE READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,499A
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/102/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEO ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-08-127-499A-11

Alignment Scores:
Pred. No.: 5,78e-117 Length: 478
Score: 2037.50 Matches: 388
Percent Similarity: 90.81% Conservative: 47
Best Local Similarity: 81.00% Mismatches: 33
Query Match: 73.50% Indels: 11
DB: 1 Gaps: 8

US-09-475-704A-3 (1-1479) x US-08-127-499A-11 (1-478)
QY 1 ATGGAGGCCCGCGCGCAAGTCTGGCGCGCGCAAGTGAAGCGCTGGAGCGCATCCGC 60
Db 1 MeGcIyAlArghAlaserValeuSerelYglYgluDeuBpArgrTgIuVsiIeArG 20
QY 61 CTGCGCCCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTG99CCAGCCGCGAG 120
Db 21 LeuArGrProGlyYglYglYslYslYsTrlyrLysLeuYshIsIleValTlPAlaserArGlu 40
QY 121 CTGGAGAAAGTTGCGCTCTGAACCCCGCGCTGCTGGAGACCAAGCGAGGGCTGMAACAATC 180
Db 41 LeuGlArGrPheAlaValAaBnProGlyLeuLeuGluTrIserGluGlyCyArGrGlnIle 60
QY 181 ATCCGCGAGGTGACCCCGCGCTGGAGACCGCGAGCGAGCGAGCTGAAGACCGCTTCAAC 240
Db 61 LeuGlYglInleuGlnProSerLeuGlnIhrGlySerGluGluDeuAaGrSerLeuTrAsn 80
QY 241 ACCGTGGCCACCTCTACTCGTGGACGAGAAAGATCGAGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTrCyArValHIsGlnArGlnIleGluIleuYsBpTrhrLySglAla 100
QY 301 CTGGACAAGTTCAGAGGAGGAGCAACAACAAGTGCCACGCAAGAAAGATCCAGACGCCGAGGCC 360
Db 101 LeuAspLyserIleGluGluGluGlnInBnYsSerLySlyAlaGlnGlnAla--Ala 119
QY 361 GCCGACAAGGCC-----AAGTGAGCGCAAGACTACCCCATCTGCGCAAGACTGCGAG 411
Db 120 AlaAspTrhclYhIsIserSerGlnValIserGlnBnTrProIleValGlnAenIleGln 139
QY 412 GGCACAGATGGTGACAGGCGCATCAAGCCCGCGACCTCGAAGCGCTGGGTGAAGGTATC 471
Db 140 GlyGlnMetValHIsGlnAlaIleSerProArGrhTrLeuAaAlaTrpValYsValVal 159
QY 472 GAGAGAAAGGCTTCAGCCCGCGAGGTGATCCCATGTTCAACGCGCTCGAGCGAGGCGCC 531
Db 160 GluGluIuYsAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluYAla 179
QY 532 ACCCCCGAGACTGAACAACATGTTGAACACCGTGGCGGGGCCACAGCGCCGCAATCAG 591
Db 180 ThrProGlnAspLeuBnSnrhMetLeuBnTrValGlyGlnHIsGlnAlaIleMetGln 199

```

```

QY 5922 ATGCTGAAGGACCATCAACGAGGAGGCGCGCATGGGACCGCGTGCACCCCGTGCAC 651
Db 200 MetLeuLySGITThrlIeAsnGluGluAAlaGluTrpMetPrGValHisProValHis 219
QY 652 GCCGGCCCCCATGCGCCCGGCAGATGCGCGACCCCGCGCGACGACATCGCGCGACC 711
Db 220 AlaGlyProIleAlaProGlyGlnMetAArgGluProArgGlySerMetPrIleAlaGlyThr 239
QY 712 ACCGACCCCTCGACGAGAGAGATCGCGTGGATGACCAAGACCCCCCATCCCGTGGGC 771
Db 240 ThrSerThrLeuGlnGluIntIleGlyTrpMetThrAsnMetProIleProValGly 259
QY 772 GACATCTCAAGCGGTGATCATCTCGGGCGTGAACAAGATCGTGGAGTGTACAGCCCC 831
Db 260 GlnIleTyTyLysArgTrpIleIleLeuGlyLeuAsnIleValAlaArgMetTySerPro 279
QY 832 GTGAGCATCTTGACATCAACGAGGCGCCCAAGAGACCCCTTCGCGCATACGTGACCGC 891
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgMetPrTyValaMetArg 299
QY 892 TTCTTCACAGACCCCTCGCGCGCGACGACAGACCCAGAGAGTGAAGATGTGAATGACCGAC 951
Db 300 PheTyTyLysThrLeuAlaArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGlu 319
QY 952 ACCCTGCTGTGTCAGACAGCCCAACCCCGCATGTGAAGACCATCTCGCGCGCTCGGCGCC 1011
Db 320 ThrLeuLeuValaGlnAsnAlaAsnProMetPrCyLysThrIleLeuLysAlaLysGlyPro 339
QY 1012 GCGCGACCTTGAGAGAGATGATGACCGCTCGCGACGCGCGTGGCGCGCCCGACCAAG 1071
Db 340 AlaAlaThrLeuGluGlnMetMetThrAlaCySGlnGlyValGlyProGlyHisLys 359
QY 1072 GCCCGCGTGTGCGCGACGCGCATGAGCCAG---GCCAAGC---AGCGTGAATGATGCAG 1125
Db 360 AlaArgValaLeuAlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
QY 1126 AAGAGCACTTCAGAGGCGCCCGCGGCGCATGTGAAGTGTCAATCGCGGACGAGAGGC 1185
Db 380 ArgGlyAsnPheArgAsnGlnArgLysMetValLysCyPheAsnCySGlyLysGluGly 399
QY 1186 CACATCGCGCGCACTGCGCGCGCCCGCGCAAGAGGCTGTGAAGTGTGGCGCAAGAG 1245
Db 400 HisThrAlaArgAsnCyPheArgAlaProArgLysGlyCyPheTrpLysGlyLysGlu 419
QY 1246 GGGACCAAGATGAAGACTGACCGACCGAGCGCCGCAACTCTCGTGGCAAGATC---TGG 1302
Db 420 GlnHisGlnMetLysAspCyThrGluArgGlnAlaAsnPheLeuGlyLysIleCyLeu 439
QY 1303 CCCAGCCCAAGAGGCGCCCGCGCAACTTCTGTGACAGGCGCCCGACGCCACCGCGCCC 1362
Db 440 ProThrArgGluGlyGln---GlyIlePhePhe-----ArgAlaAsnGlnSerGlnGln 456
QY 1363 CCGCGCGAGAGCTCGCTTCGTGAGAGACC---ACCGCGCGCGCAAGACGAGAGACC 1416
Db 457 ProHisIlePhePheArgAlaAsnGlnSerGlnGlnProHisGlnLysArgAlaSer 475

```

```

: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/482,847
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/127,499
: FILING DATE: 28-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 51916/104/INBI
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 478 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
:
: US-08-482-847-11
:
: Alignment Scores:
: Pred. No.: 5,78e-117 Length: 478
: Score: 2037.50 Matches: 388
: Percent Similarity: 90.81% Conservative: 47
: Best local Similarity: 81.00% Mismatches: 33
: Query Match: 73.50% Indels: 11
: DB: 1 Gaps: 8
:
: US-09-475-704A-3 (1-1479) x US-08-482-847-11 (1-478)
:
: Oy 1 ATGGAGCGCCCGCGCGCAAGCTCTGGCGCGGCGGCAAGTGAAGCGCTGGAGCGCATCCGC 60
: Db 1 MeGcIYhIAIhrgAlaservAlleuervglYglgluIuabpaGrTgIuYsIleIhrg 20
: Oy 61 CTGCGCGCCCGCGCGCAAGAGTGTCAATGAGAGCAAGCACTGTGTGTGGCCAGCGCGAG 120
: Db 21 LeuAtrPrGclYglYsIeYsIySrlYrIySleuYshIstIeValTTPAlaserArgIu 40
: Oy 121 CTGGAAGAATTGGCCCTGAACCCCGGCGCTGTGTGAAGACACGAGGCGTGTGACAGATC 180
: Db 41 LeuGlIuArPhelAlaValaAnPrGclYleuSrlYleuGlIuTrSerGlugIuYcyarGInIle 60
: Oy 181 ATCCGCGAGGTGAACCCCGCCCTGTGAAGACCGGCGAGGAGCTGGAAGAGCTGTTCAC 240
: Db 61 LeuGlYglIneudIuProserIeuGlInThrgIySerGlugIuIuAuYserLeuYrAsn 80
: Oy 241 ACCGTGAGCCACCTGTATGTGCTGTGACGAGAGATCGAGTCCGCGACCAACAGAGAGCC 300
: Db 81 ThValAlaThIreutYrCySValhIseIuArglIegIuIleYsAepThIuYsGlua 1000
: Oy 301 CTGAGCAAGATCGAGGAGGAGGAGCAACAAGTGTCCAGCAAGATTCAGAGGCGGAGGCC 3606
: Db 101 LeuAspIyIeIedIuclIuclInuIuInuYsSerIySrlYsIuYsAlaGlInIaIe--Ala 1199
: Oy 361 GCCGACAAAGGC-----AAGTGAGCCAGAACTACCCCATCTGTGCAAGACTTCAG 4111
: Db 120 AlaAspThGlYhIaserIerInValSerGlInuYrProIleValIglInAsrIleGlh 1399
: Oy 412 GGCAGATGAGTGTGACCAAGGCGATCAAGCCCGGACCCGTGAAGCGCTGTGGGGAAGGTATC 4717
: Db 140 GlyIuIuIeValhIseIuAlaIleSerIroArghIuIeuaMAlaIrrPaIlyValVal 1599
: Oy 472 GAGAGAAAGGCTTACGCCCGAGGATGATCCCATGTTCACCGGCTGTAGCGAGGCGCC 5311

```



```
Db      140 GlyMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIlySerValVal 159
QY      472 GAGGAGAAAGGCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCC 531
Db      160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluIlyAla 179
QY      532 ACCCCGAGGACCTGAAACGATGTTGAACACCGTGGGGGCCACGAGCGCGCATGACG 591
Db      180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGln 199
QY      592 ATGCTGAAGAGACACCATCAACGAGAGCGCGCGAGTGGGACCGCGTGACCCCGTGAC 651
Db      200 MetLeuLysGluThrIleAsnGluGlnAlaIleGluTrpAspArgValHisProValHis 219
QY      652 GCGGCGCCCATGCGCCCCGCGGCGAGATGCGGAGCGCGCGGCGAGCATGCGCGGACCC 711
Db      220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239
QY      712 ACCAGACCCCTGAGAGAGAGATGCGCTGATGACCAAGCAACCCCGCATCCCGTGGGC 771
Db      240 ThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProIleProValGly 259
QY      772 GACATCTAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGGATGTACAGCCCG 831
Db      260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysAlleValArgMetTyrSerPro 279
QY      832 GTGAGCATCTGAGCATCAAGCAGGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGC 891
Db      280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY      892 TTCTTCAAGACCTGCGCGCGCGAGCAGACCCAGAGAGTGAAGACTGATGACCGAC 951
Db      300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnIleValLysAsnTrpMetThrGln 319
QY      952 ACCCTGCTGTGAGAAACGCCAACCCGAGCTGCAAGACCATCTGCGCGCTCTCGGCGCC 1011
Db      320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY      1012 GCGCGCAGCTGAGAGAGATGATGACCGCTGCGAGGGCGTGGCGCGCCCGACCAAG 1071
Db      340 AlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY      1072 GCGCGGTGCTGGCGGAGCGGATGAGCCAGGCCAACCAACGAC-----GTGATGATGACG 1125
Db      360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGln 379
QY      1126 AAGAGCAACTTCAAGGGGCGCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGAGGGC 1185
Db      380 ArgGlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGlnGly 399
QY      1186 CACATGCGCGCAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGAAATGCGGCAAGAG 1245
Db      400 HisThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGln 419
QY      1246 GGGCAACCAATGAAGATGCAACCGAGCGCGCAGGCCCACTTCTTGGGCAAGATC 1299
Db      420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIle 437
```

Search completed: September 1, 2005, 23:19:32  
Job time : 47.739 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 1, 2005, 22:09:18 ; Search time 378.412 Seconds  
(without alignments)  
3078.466 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 2772  
Sequence: 1 atggggcggccgcgcagcat.....acgacccctgacgcagcraa 1479

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 3548624

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=x1p  
-Q=/cg2\_1/USPTO.spool.p/US09475704/runat.01092005.160752.19548/app.query.fasta\_1.3342  
-DB=Published Applications AA -QMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62  
-TRANS=human0.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100  
-THR.MIN=0 -ALLEN=15 -MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09475704 @CGN 1.1 233 @runat.01092005.160752.19548  
-NCPU=6 -ICPU=3 -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSBIOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:\*

1: /cg2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cg2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cg2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cg2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cg2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
6: /cg2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
7: /cg2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cg2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cg2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cg2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cg2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cg2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cg2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cg2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cg2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cg2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cg2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep:\*  
18: /cg2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
19: /cg2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep:\*  
20: /cg2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
21: /cg2\_6/ptodata/1/pubpaa/US6\_NEW\_PUB.pep:\*  
22: /cg2\_6/ptodata/1/pubpaa/US6\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2606	94.0	492	10	US-09-899-575-17 Sequence 17, Appl
2	2464	88.9	492	18	US-10-780-507-96 Sequence 96, Appl
3	2462.5	88.8	493	18	US-10-780-507-98 Sequence 98, Appl
4	2461	88.8	492	18	US-10-780-507-97 Sequence 97, Appl
5	2448	88.3	492	16	US-10-332-413-2 Sequence 2, Appl
6	2448	88.3	492	18	US-10-494-131-16 Sequence 36, Appl
7	2435	87.8	492	9	US-09-991-258-5 Sequence 5, Appl
8	2435	87.8	492	18	US-10-929-234-5 Sequence 8, Appl
9	2435	87.8	631	16	US-10-332-413-8 Sequence 110, App
10	2423	87.4	492	14	US-10-339-217-110 Sequence 109, App
11	2416	87.2	492	14	US-10-339-217-109 Sequence 140, App
12	2415.5	87.1	491	14	US-10-339-217-140 Sequence 104, App
13	2407.5	86.9	491	14	US-10-339-217-104 Sequence 143, App
14	2396.5	86.5	491	14	US-10-339-217-143 Sequence 108, App
15	2382	85.9	496	14	US-10-339-217-108 Sequence 106, App
16	2381.5	85.9	487	14	US-10-339-217-106 Sequence 107, App
17	2381.5	85.9	497	14	US-10-339-217-107 Sequence 105, App
18	2344	84.6	508	14	US-10-339-217-105 Sequence 22, Appl
19	2343	84.5	502	10	US-09-899-575-22 Sequence 126, App
20	2328.5	84.0	494	14	US-10-339-217-126 Sequence 18, Appl
21	2287	82.9	500	14	US-10-339-217-18 Sequence 47, Appl
22	2287	82.9	500	18	US-10-780-507-47 Sequence 49, Appl
23	2281	82.3	500	18	US-10-780-507-49 Sequence 22, Appl
24	2277	82.1	498	16	US-10-325-468-22 Sequence 8, Appl
25	2266	81.7	500	16	US-10-325-468-8 Sequence 34, Appl
26	2263	81.6	500	16	US-10-325-468-14 Sequence 27, Appl
27	2262	81.6	498	16	US-10-325-468-27 Sequence 48, Appl
28	2259	81.5	500	18	US-10-780-507-48 Sequence 121, App
29	2258	81.5	500	14	US-10-339-217-121 Sequence 85, Appl
30	2256	81.4	500	14	US-10-059-271-85 Sequence 4, Appl
31	2256	81.4	500	14	US-10-842-652-4 Sequence 6, Appl
32	2256	81.4	500	17	US-10-842-658-6 Sequence 1, Appl
33	2252.5	81.3	499	15	US-10-286-734-1 Sequence 17, Appl
34	2252.5	81.2	500	16	US-10-325-468-17 Sequence 26, Appl
35	2251	81.2	500	9	US-09-968-355-26 Sequence 17, Appl
37	2251	81.2	583	9	US-09-968-355-17 Sequence 42, Appl
38	2249	81.1	500	14	US-10-223-172A-42 Sequence 3483, App
39	2249	81.1	500	14	US-10-224-999A-3483 Sequence 136, App
40	2246.5	81.0	501	14	US-10-339-217-136 Sequence 35, Appl
41	2245	81.0	1350	16	US-10-952-060-35 Sequence 35, Appl
42	2245	81.0	1350	16	US-10-380-641-35 Sequence 35, Appl
43	2245	81.0	1350	17	US-10-636-730-35 Sequence 144, App
44	2237	80.7	557	14	US-10-339-217-144 Sequence 8, Appl
45	2235	80.6	512	17	US-10-634-165-8

## ALIGNMENTS

RESULT 1  
US-09-899-575-17  
; Sequence 17, Application US/09899575  
; Publication No. US2003023961A1  
; GENERAL INFORMATION:  
; APPLICANT: Zur Megele, Jan  
; APPLICANT: Barnett, Susan W.  
; APPLICANT: Egnelbrecht, Susan  
; APPLICANT: van Rensburg, Estrelita Janse  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: PP01631.102  
; CURRENT APPLICATION NUMBER: US/09/899,575  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 09/475,704  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 492

```

; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-899-575-17

Alignment Scores:
Pred. No.: 3,19e-129 Length: 492
Score: 2606.00 Matches: 492
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.01% Indels: 0
DB: 10 Gaps: 0

US-09-475-704a-3 (1-1479) x US-09-899-575-17 (1-492)

QY 1 ATGGGCGCCGCGCCGATCCTGCGGCGCGCAACCTGGAAGCGCTGGAGCGCATCCGC 60
   |||||
Db 1 MeGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspAlaTrpGluArgIleArg 20

QY 61 CTGCGCCCGCGGCAAGAGTGTCAATGATGAGACCTGGTGTGGCCAGCCGCGAG 120
   |||||
Db 21 LeuArgProGlyGlyLeuLeuSerGlyTrpMetMetLeuValTrpAlaSerArgGlu 40

QY 121 CTGGAAGAATTGCGCCTGAACCCGCGCTGTGTGAGACCAAGGCGCTGCAAGATC 180
   |||||
Db 41 LeuGluLeuPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysIysGlnIle 60

QY 181 ATCCGCGCAGCTGACCCCGCCCTGACAGCCGCGAGGAGGAGCTGAAGCGCTTCAAC 240
   |||||
Db 61 IleArgGlnLeuHisProAlaLeuGlnThrGlySerGluGluLeuLeuSerLeuPheAsn 80

QY 241 ACCGTGGCCACCTGTACTGCGTGACAGAGATCAGAGTCCGCGACCAAGAGAGCC 300
   |||||
Db 81 ThrValAlaThrLeuLeuTrpCysValHisGluIleValArgAspThrIysGlnAla 100

QY 301 CTGGAACAATGAGAGAGAGAGCAACAAGTCCACAGACCAAGATCCAGAGCGCGAGCC 360
   |||||
Db 101 LeuAspLysIleGluGluGlnGluGlnAsnLysCysGlnGlnLysIleGlnGlnAlaGluAla 120

QY 361 GCGCAAGAGGAGGAGTGAGCAAGACTACCCATGTCGACAGACCTGAGGCGCGAGT 420
   |||||
Db 121 AlaAspLysGlyLeuValSerGlnAsnTrpProIleValGlnAsnLeuGlnGlnMet 140

QY 421 GTGCAACGAGCGCATCAGCCCGCGACACCTGGAAGCTGATGAGAGAGAG 480
   |||||
Db 141 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysAlaIleGluGluLys 160

QY 481 GCGTTAGCGCCGAGGTGATCCCATGTTCAACCGCTTGAAGAGGCGCGCACCCCGAG 540
   |||||
Db 161 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 180

QY 541 GACCTGAACAGATGTTGAACAACCGTGGGCGGCAAGCGCGCATGGAATGCTGAG 600
   |||||
Db 181 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 200

QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCTGACCGCGCGCC 660
   |||||
Db 201 AspThrIleAsnGlnGluAlaIleGluTrpAspArgValHisProValHisAlaGlyPro 220

QY 661 ATGCGCCCGCGCGAGATGCGCGAGAGCCCGCGGAGGAGATGCGCGGCAACCAAGACC 720
   |||||
Db 221 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerTrp 240

QY 721 CTGCAAGGAGAGATGCGCTGATGACCAAGACCCCGCATCCCTGGGCGAGACTTAC 780
   |||||
Db 241 LeuGlnGlnGlnIleAlaTrpMetTrpSerAsnProProIleProValGlyAspIleTrp 260

QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATGTCGATGATGACGCGCGGTGAGATC 840
   |||||
Db 261 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProValSerIle 280

QY 841 CTGGAATCAACGAGGCGCCCAAGAGAGCCCTTCGGAATCACTGAGACGCTTCTTCAAG 900
   |||||
Db 281 LeuAspIleLeuGlnGlnIleProLysGluProPheArgAspTrpValAspArgPhePheLys 300
```

```

QY 901 ACCGTGGCGCGCGAGAGAGACCCAGAGAGGTGAAGAATGAGATGACCGACACCTGCTG 960
   |||||
Db 301 ThrLeuArgAlaGlnGlnSerThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu 320

QY 961 GTGCAAAAGCCCAACCCGAGATGCAAGACATCTCTGCGCGCTTGGCCCGCGCGCAGC 1020
   |||||
Db 321 ValGlnAsnAlaAsnProAspCysIleLeuArgIleLeuArgAlaLeuGlyProGlyAlaSer 340

QY 1021 CTGGAAGATGATGATACCGCTGCGGAGGCGGTGGCGGCGCCCAAGCAAGGCGCGTG 1080
   |||||
Db 341 LeuGluGluMetMetThrAlaCysGlnGlnIleValGlyGlyProSerHisLeuValArgVal 360

QY 1081 CTGCGCGAGGAGATGAGCAGGCGCAACCAAGCGGTGATGATGAGCAAGAGCAACTTCAAG 1140
   |||||
Db 361 LeuAlaGluAlaMetSerGlnAlaAsnThrSerValMetMetGlnLysSerAsnLys 380

QY 1141 GCGCCCGCGCGCATGCTCAAGTGTCTTCAACTGCGCGCAAGAGGCGCACATGCGCCGCAAC 1200
   |||||
Db 381 GlyProArgArgIleValIleCysPheAsnGlySerGlyGluGlnHisIleAlaArgAsn 400

QY 1201 TGCCCGCGCGCGCGCAAGAGGCGCTGTGGAAGTGTGCGCAAGAGGCGCAACATGTAAG 1260
   |||||
Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlnHisIleGlnMetLys 420

QY 1261 GACTGACAGGAGCGCGCAGCGCAACTTCTGGGCAAGATCTGGCCGCGCACAGGCGCGC 1320
   |||||
Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440

QY 1321 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCGCGCGCGCGCGCGAGAGCTTCCGC 1380
   |||||
Db 441 ProGlyAsnPheLeuGlnSerArgProGluProThrAlaProProAlaGluSerPheArg 460

QY 1381 TTCGAGAGAGACCAACCCCGCGCAAGAGAGAGAGAGACCGGAGACCTTGACCAAGC 1440
   |||||
Db 461 PheGluGluThrThrProGlyGlnLysGlnLysSerLysAspArgGluThrLeuThrSer 480

QY 1441 CTGAAGAGCTGTTGGCAAGACCCCGTGAAGCCAG 1476
   |||||
Db 481 LeuLysSerLeuPheGlyAsnAspProLeuSerGln 492

RESULT 2
US-10-780-507-96
; Sequence 96, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MOLLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Most recent common ancestor reconstruction of clade C gag protein
; OTHER INFORMATION: sequence
```

US-10-780-507-96

## Alignment Scores:

Pred. No.:	9,5e-122	Length:	492
Score:	2464.00	Matches:	464
Percent Similarity:	97.57%	Conservative:	17
Best Local Similarity:	94.12%	Mismatches:	10
Query Match:	88.89%	Indels:	2
DB:	18	Gaps:	2

US-09-475-704A-3 (1-1479) x US-10-780-507-96 (1-492)

```

QY 1 ATGGGCGCGCGCGCGCATCTCTGCGGCGGCAAGCTGGAACGCTGGAGCGCATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspThrTrpGluysIleArg 20
QY 61 CTGGCGCGCGCGCGCAAGAGTGCTACATGATGAGACACCTGGTGTGGCCAGCCGCGAG 120
DB 21 LeuArgProGlyGlyysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 40
QY 121 CTGGAGAAATGTCGCTGAAACCCCGGCTGCTGGAGAACAGGAGGCTGGACAGCATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysIysGlnIle 60
QY 181 ATCCGCGCAGCTGCAACCCCGCTGACAGCCGCGGAGAGAGCTGAAGAGCTGTTCAAC 240
DB 61 IleGlyGlnIleuGlnProAlaLeuGlnThrGlyThrGluGluLeuysSerLeuysIleuys 80
QY 241 ACCGTGGCAACCTGTAAGCTGCTGCAAGAGATGAGAGTCCGCGACCAAGAGGCC 300
DB 81 ThrValAlaThrLeuLeuysCysValHisGlnArgIleGluValArgAspThrIleuysIle 100
QY 301 CTGGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 101 LeuAspLeuysIleGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 120
QY 361 GCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 121 AlaAsp---GlyLeuValSerGlnAsnIleuysIleuysIleuysIleuysIleuysIleuys 139
QY 421 GTGGCAACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaIleuysIleuysIleuysIleuys 159
QY 481 GCGTTACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluValIleuysIleuys 179
QY 541 GACGTGAACAGATGTGAACACCGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleuysIleuysIleuys 199
QY 601 GACACCATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 200 AspThrIleAsnGlnIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 219
QY 661 ATCCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 220 ValAlaProGlyGlnMetArgIleuysIleuysIleuysIleuysIleuysIleuysIleuys 239
QY 721 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 240 LeuGlnIleuGlnIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 259
QY 781 AAGCGGTGATCATCTGGGCTGGAACAAGATCTGGAGATGATGAGAGAGAGAGAGAGAG 840
DB 260 LysArgTrpIleIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 279
QY 841 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 280 LeuAspIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 299
QY 901 ACCGTGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

```

```

DB 300 ThrLeuArgAlaGlnIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 319
QY 961 GTGGAGAAACGCAACCCCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 320 ValGlnAsnAlaAsnProAspIleuysIleuysIleuysIleuysIleuysIleuysIleuys 339
QY 1021 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 340 LeuGlnIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 359
QY 1081 CTGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
DB 360 LeuAlaGlnAlaMetSerGlnAlaAsnThrAsnIleuysIleuysIleuysIleuysIleuys 379
QY 1138 AAGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
DB 380 LysGlyProArgArgIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 399
QY 1198 AACTGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
DB 400 AsnCysArgAlaProArgIleuysIleuysIleuysIleuysIleuysIleuysIleuys 419
QY 1258 AAGAGTGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
DB 420 LysAspCysThrGlnArgGlnAlaAsnIleuysIleuysIleuysIleuysIleuysIleuys 439
QY 1318 CGCGCGCGCAACTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
DB 440 ArgProGlyAsnPheLeuGlnIleuysIleuysIleuysIleuysIleuysIleuysIleuys 459
QY 1378 CGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
DB 460 ArgPheGlnIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuysIleuys 479
QY 1438 AGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
DB 480 SerLeuysSerLeuysPheGlySerAspProLeuysIleuysIleuysIleuysIleuys 492

```

## RESULT 3

```

US-10-780-507-98
; Sequence 98, Application US/10780507
; Publication No. US2005013787A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10780,507
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Minimum of means center of tree reconstruction of clade C gag pro
US-10-780-507-98

```

## Alignment Scores:

Pred. No.:	1,14e-121	Length:	493
Score:	2462.50	Matches:	461
Percent Similarity:	97.36%	Conservative:	19
Best Local Similarity:	93.51%	Mismatches:	12
Query Match:	88.83%	Indels:	1
DB:	18	Gaps:	1

US-09-475-704a-3 (1.1479) x US-10-780-507-98 (1.493)

```

QY 1 ATGGGCGCCCGCCGACATCTCGCGCGCGGACAGCTGAGCGCTGGAGCGCATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspThrTrpIleArg 20
QY 61 CTGGCGCCCGCGGAGAGAGTGTACTGATGAGACCTGTGTGGCCGAGCGCGAG 120
DB 21 LeuArgProGlyGlyLeuValshstYrMetLeuValshLeuValTrpAlaSerArgGln 40
QY 121 CTGAGAAAGTTGCGCTGAAACCCGCGCTGTGAGACCGAGCGAGGGCTGCAAGCATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGlnGlyCysIysGlnIle 60
QY 181 ATCCGCGACCTGACACCCGCGCTGCAGACCGGACGAGAGAGCTGAAAGCTGTTCAC 240
DB 61 MetYsGlnLeuGlnProAlaLeuGlnThrGlyThrGlnGluLeuArgSerLeuYrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGCGACGAGAAAGTGAAGTCCGCGACCAAGAGGCG 300
DB 81 ThrValAlaThrLeuYrCysValshGlnYsIleGlnValArgAspThrYsGlnAla 100
QY 301 CTGGAACAGATCGAGAGAGAGAGAGACAAAGTGCACAGACAGAGATCCAGAGCGCGAGCC 360
DB 101 LeuAspYsIleGlnGlnGlnIleAsnYsSerGlnGlnIleThrGlnIleAlaGlnAla 120
QY 361 GCGGACAGAGGCGAAGGTGACCGCAAGACTACCCATGTGTCAGAACTCGAGCGCGCATG 420
DB 121 AlaAlaAspGlyLeuValSerGlnAsnYrProIleValGlnAsnLeuGlnGlnIleMet 140
QY 421 GTGACCAAGGCGCATAGCCCGCGACCCCTGAAAGCGCTGTGTAAGTATCGAGAGAG 480
DB 141 ValshGlnAlaIleSerProArgThrLeuAsnAlaTrpValYsAlaIleGlnGlnYs 160
QY 481 GCGTTACGCGCGAGGTGATCCCATGTTACCGCGCTGAGCGAGCGCGCGACCCCGCAG 540
DB 161 AlaPheSerProGlnValIleProMetPheThrAlaLeuSerGlnGlyAlaThrProGln 180
QY 541 GACCTGAACAGATGTTGAACAACCGTGGCGCGGCACACGCGCGCATGCAATGCTGAAG 600
DB 181 AspLeuAsnThrMetLeuAsnThrValGlyGlnshGlnAlaAlaMetGlnMetLeuYs 200
QY 601 GACACCATCAAGAGAGCGCGCGAGTGGAGCGCGGTGCACCGCTGACGCGCGCGCC 660
DB 201 AspThrIleAsnGlnIleAlaIleGlnIleAspArgLeuIleAspProAlaIshIleGlyPro 220
QY 661 ATCGCGCGCGCGAGTGGCGAGCGCGCGCGAGCGAGCATGCGCGGACCAACGACGACC 720
DB 221 ValAlaProGlyGlnMetArgGlnProArgGlySerAspIleAlaGlyThrThrSerThr 240
QY 721 CTGCGAGAGAGATGCGCTGATGACAGCAACCCCGCATCCCGCTGGGGCAATCTAC 780
DB 241 LeuGlnGlnGlnIleAlaIleTrpMetThrSerAsnProProValProValGlyAspIleYr 260
QY 781 AAGCGGTGATCATCTGGGCGTGAACAGATCGTGGGATGTATACAGCCCGCTGAGCANC 840
DB 261 YsAspArgTrpIleIleLeuGlyLeuAsnYsIleValArgMetYrSerProValSerIle 280
QY 841 CTGGAATCAAGAGGCGCGCAAGAGAGCGCTTCGCGACTAGTGGACCGCTTCTTCAAG 900
DB 281 LeuAspIleIleYsGlnGlyProYsGlnProPheArgAspYrValAspArgPheYs 300
QY 901 ACCCTGGCGCGGAGAGAGACCCGAGAGTGAAGAACTGATGACGACCACTGTGCTG 960
DB 301 ThrLeuArgAlaGlnGlnAlaThrGlnAspValYsAsnTrpMetThrAspThrLeuLeu 320

```

```

QY 961 GTGGAAGCGCCAAACCCGAGCTGCAAGAGCAATCCCTGGCGGCTTGCGGCCGCGGCGCAGC 1020
DB 321 ValGlnAsnAlaAsnProAspCysLeuYrThrIleLeuArgAlaLeuGlyProGlyAlaThr 340
QY 1021 CTGAGAGATGATGACCGCTGCGGAGCGGTGGCGCGCGCCAGCACAAGCGCGCGCTG 1080
DB 341 LeuGlnGluMetMetThrAlaCysGlnGlnIleValGlyGlyProGlnYshIleYsAlaArgVal 360
QY 1081 CTGGCGGAGGCGATGAGCGCAGGCC---AAGCAACCGTGAATGATGACAAAGCAACTTC 1137
DB 361 LeuAlaGlnAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgSerAsnPhe 380
QY 1138 AAGGCGCGCGCGCATGTGTCAGAGCTTCAACTGCGCGGCAAGAGAGGCGCACATCGCGCG 1197
DB 381 YsGlyProYsArgGlyIleValYsCysPheAsnYsGlyYsGlnGlnIleAlaArg 400
QY 1198 AACTCGCGCGCGCGCGGAGAGAGCGCTGTGAAGTGCAGGAGCGCAAGAGGCGCACAGATG 1257
DB 401 AsnYsArgAlaProArgYsLeuGlyCysThrYsCysGlyYsGlnGlnIleMet 420
QY 1258 AAGACTGACCGGAGCGCGCAGCGCAACTTCTGGGCAAGATCTGGCCGACCAAGCGC 1317
DB 421 YsAspCysThrGlnArgGlnAlaAsnPheLeuGlyYsIleTrpProSerHisYsGly 440
QY 1318 CGCGCGCGCACTTCTGTCAGAGCGCGCGGAGCGCACCGCGCGCGCGGAGAGCTTC 1377
DB 441 ArgProGlyAsnPheLeuGlnSerArgProGlnProThrAlaProProIleGlnSerPhe 460
QY 1378 CGCTTCGAGAGACCAACCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
DB 461 ArgPheGlnGlnIleThrThrProAlaProYsGlnGlnIleProYsAspArgGlnProLeuThr 480
QY 1438 AGCTGAAGAGCTGTGTGCGCAAGACCCCTGAGCGCAG 1476
DB 481 SerLeuYsSerLeuPheGlySerAspProLeuSerGln 493

RESULT 4
US-10-780-507-97
; Sequence 97, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Least squares center of tree reconstruction of clade C gag protei
US-10-780-507-97

Alignment Scores:
Pred. No.: 1.37e-121 Length: 492

```

Score: 2461.00 Matches: 462  
 Percent Similarity: 97.57% Conservative: 19  
 Best Local Similarity: 93.71% Mismatches: 10  
 Query Match: 88.78% Indels: 2  
 DB: 18 Gaps: 2

US-09-475-704a-3 (1-1479) x US-10-780-507-97 (1-492)

QY 1 ATGGGCGCCCGGCGAGCATCTCGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCG 60  
 DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspThrTrpAlaSerArgGlu 20  
 QY 61 CTGGCGCCCGGCGGCAAGAGTCTACATGATGAACCACTGTGTGGCCGCGCGG 120  
 DB 21 LeuArgProGlyGlyLeuYshIstYrMetLeuYshIleValTrpAlaSerArgGlu 40  
 QY 121 CTGAGAAAGTTGCGCTGAAACCCGCGCTGTGAGACCGAGCGAGGCTGCAAGCATC 180  
 DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysIleGlnIle 60  
 QY 181 ATCCGCGACCTGACCCCGCTGACAGCCGCGACGAGAGAGCTGAAGCTGTTCAC 240  
 DB 61 MetYsgInLeuGluProAlaLeuGluThrGlyThrGluGluLeuArgSerLeuYrAsn 80  
 QY 241 ACCGTCGACCTGTACTGCTGTCACGAGAAAGTGAAGTCCGCGACCAAGAGCGC 300  
 DB 81 ThrValAlaThrLeuYrCysValHisGluYshIleGluValArgAspThrIysGluAla 100  
 QY 301 CTGCAACAAGATCGAGAGGAGCAACAAGTGCACAGACAAGATCCAGAGCGCGAGCC 360  
 DB 101 LeuAspYshIleGluGluGluGlnAsnYshSerGlnGluYshThrGlnIleGluAla 120  
 QY 361 GCCGACAAAGGCGAAGGTGAAGCAAACTACCCCATCTGTGACAGACTGTGAGGCGCAGATG 420  
 DB 121 AlaAsp---GlyValSerGlnAsnYrProIleValGlnAsnLeuGlnGlyGlnMet 139  
 QY 421 GTGACCAAGGCGATACGCCCGCGACCTCGAAGCGCTGGTGAAGTATCGAGAGAAAG 480  
 DB 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValHisIleGluGluYsh 159  
 QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTCACCGCCCTGAGGAGGCGCGACCCCGCAG 540  
 DB 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 179  
 QY 541 GACTGAACACGATGTTGAACACCGTGGCGGCGCACAGCGCGCATGCGATGCTGAAG 600  
 DB 180 AspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuYsh 199  
 QY 601 GACACATCAAGAGAGGCGCGCATGTGGACCGCGTGCACCCGCTGACCGCGCGCCG 660  
 DB 200 AspThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisIleAlaGlyPro 219  
 QY 661 ATCGCCCGCGGAGATGCGCGAGCGCGCGGCGAGGACATGCGCGGCGACCAAGACC 720  
 DB 220 ValAlaProGlyGlnMetArgGluProArgGlySerPheIleAlaGlyThrThrSerThr 239  
 QY 721 CTGCAAGAGACGATGCTGATGACCAACACCCCGCATCCCGTGGCGGACATCTAC 780  
 DB 240 LeuGlnGluGlnIleAlaTrpMetThrSerAsnProProValProValGlyAspIleYr 259  
 QY 781 AAGCGGTGATGATCTCGGCTGGAACAAGATCGTCGGAATTAAGCCCGCTGAGCATC 840  
 DB 260 LysArgTrpIleIleLeuGlyLeuAsnYshIleValArgMetYrSerProValSerIle 279  
 QY 841 CTGCAATCAAGACGAGCGCGCAAGAGCGCTTCGCGCATTAAGTGAAGCGCTTCTTCAAG 900  
 DB 280 LeuAspIleIleYshGlnGlyProYshGluProPheArgAspYrValAspArgPhePheYsh 299  
 QY 901 ACCCTGCGCGCGAGAGCAAGCAACCAAGAGGTGAAGAACTGTATGACCAAGCATCTGCTG 960  
 DB 300 ThrLeuArgAlaGlyGlnIleAlaThrGlnAspValYshAsnTrpMetThrAspThrLeuLeu 319  
 QY 961 GTGCAAGACGCCAAGCCCGCATGCAAGACCATCTCGCGGCTCTGCGGCCGCGCGCACG 1020

DB 320 ValGlnAsnAlaAsnProAspYshYshThrIleLeuArgAlaLeuGlyProGlyAlaThr 339  
 QY 1021 CTGAGAGATGATGATACCGCTGCGAGGCGCTGTGGCGCGCCCGACCAAGAGCGCGCTG 1080  
 DB 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisIleYshAlaArgAla 359  
 QY 1081 CTGGCCGAGGCGATGAGCCAGGCG---AACAACAGCTGATGATGACAGAAAGCAATTC 1137  
 DB 360 LeuAlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgSerAsnPhe 379  
 QY 1138 AAGGCGCCCGCGCATGTGCAAGTCTTCAATCGCGGCAAGAGGCGCGACATGCGCGCG 1197  
 DB 380 LysGlyProYshArgGlyValYshCysPheAsnYshGlyYshGluGlyHisIleAlaArg 399  
 QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGTGAAAGTGTGCGCAAGAGAGGCGCACAGATG 1257  
 DB 400 AsnYshArgAlaProArgGlyYshGlySerTrpYshCysGlyYshGluGlyHisIleGlnMet 419  
 QY 1258 AAGACTGACCGAGCGCGCGACGCAACTTCTGTGGCAAGATCTGGCCGACCAAGGCGC 1317  
 DB 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyYshIleTrpProSerHisIleYshGly 439  
 QY 1318 CGCCCGCGCACTTCTGTCAGAGCGCGCGCGACCGCGCGCGCGCGCGAGAGCTTC 1377  
 DB 440 ArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProProAlaGluSerPhe 459  
 QY 1378 CGCTTCGAGAGACCAACCCCGCGCGCAAGAGAGAGCAAGACCGCGAGACCTTGACC 1437  
 DB 460 ArgPheGluGluThrThrProAlaProYshGlnGluProYshAspArgGluProLeuThr 479  
 QY 1438 AGCTGAAGACCTGTTCGCGCAACGACCCCTTGAGCCAG 1476  
 DB 480 SerLeuYshSerLeuPheGlySerAspProLeuSerGln 492

RESULT 5  
 US-10-332-413-2  
 / Sequence 2, Application US/10332413  
 / Publication No. US20040116660A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Swanson, Robert Edward  
 / APPLICANT: Morris, Lynn  
 / APPLICANT: Karim, Salim Abdool  
 / APPLICANT: Williamson, Carolyn  
 / TITLE OF INVENTION: Process for the Selection of HIV-1 Subtype C Isolates, Selected H  
 / FILE REFERENCE: 45669-281993  
 / CURRENT APPLICATION NUMBER: US/10/332,413  
 / PRIOR FILING DATE: 2003-09-12  
 / PRIOR APPLICATION NUMBER: PCT/IB01/01208  
 / PRIOR FILING DATE: 2001-07-09  
 / PRIOR APPLICATION NUMBER: US 60/216,995  
 / PRIOR FILING DATE: 2000-07-07  
 / PRIOR APPLICATION NUMBER: ZA 2000/3437  
 / PRIOR FILING DATE: 2000-07-10  
 / PRIOR APPLICATION NUMBER: ZA 2000/4924  
 / PRIOR FILING DATE: 2000-09-15  
 / NUMBER OF SEQ ID NOS: 32  
 / SOFTWARE: PatentIn version 3.1  
 / SEQ ID NO 2  
 / LENGTH: 492  
 / TYPE: PRT  
 / ORGANISM: Human immunodeficiency virus type 1  
 / US-10-332-413-2

Alignment Scores:  
 Pred. No.: 6,61e-121 Length: 492  
 Score: 2448.00 Matches: 461  
 Percent Similarity: 97.36% Conservative: 19  
 Best Local Similarity: 93.51% Mismatches: 11  
 Query Match: 88.31% Indels: 2  
 DB: 16 Gaps: 2

US-09-475-704a-3 (1-1479) x US-10-332-413-2 (1-492)

```

QY 1 ATGGGGGGGGGGGGGCGGATCTTGGCGGGGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspLysTrpGlyLysIleArg 20
QY 61 CTGGCGGGGGGGGCAAGAGTGTCTACATGATGAACCACTGGGTGGGCGAGCGCGGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTrpMetLeuLysHisIleValTrpAlaSerArgGly 40
QY 121 CTGGAGAGATTGGCTTGAACCCCGGCTGTGGAGACCAAGCGGGCTGCAACAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLeuGlnIle 60
QY 181 ATCCGCGCACTGACCCCGGCTGGAGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTrpAsn 80
QY 241 ACCGTGGCAACCTTACTGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 81 ThrValAlaThrLeuTrpCysValHisGlyLysIleGlyValArgAspThrLysGlyAla 100
QY 301 CTGGACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 101 LeuAspLysIleGlyGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 120
QY 361 GCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 121 AlaAsp--GlyLysValSerGlnAsnTrpProIleValGlnAsnLeuGlnGlyMet 139
QY 421 GTGCACAGAGGCGATCAGCCCGGCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 159
QY 481 GCGTTGAGCCCGGAGGTGATCCCATGTTTCAACCGCTTGAAGCGGAGCGGCAACCCCG 540
Db 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 179
QY 541 GACCTGAACAGCATGTTGAACAACCGTGGCGGCGGCAACAGCGGCGGCGGCGGCGG 600
Db 180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 199
QY 601 GACACCATCAAGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 200 AspThrIleAsnGluGluAlaIleGluTTPAspArgValHisProValHisAlaGlyPro 219
QY 661 ATGCGCCCGGCGGAGTGGCGGAGCGCGGCGGAGAGATGCGCGGCGGCAACAGAGAG 720
Db 220 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGGAGAGAGAGATGCGCTGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 240 LeuGlnGluGlnIleAlaTrpMetTrpSerAsnProIleProValIleLysAspIleTrp 259
QY 781 AAGCGGTGATCATCTGGGCGCTGAAACAAGATGTGCGATGACAGCGCGGAGAGATC 840
Db 260 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProValSerIle 279
QY 841 CTGGACATCAAGAGAGAGCGGCGGAGAGAGCGGCGGAGATGAGCGGCTTTCTTCAAG 900
Db 280 LeuAspIleArgGlnGlyProLysGluProPheArgAspTrpValAspArgPheLys 299
QY 901 ACCCTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 300 ThrLeuArgAlaGluGlnIleAlaThrGlnGluValLysAsnTrpMetThrAspThrLeu 319
QY 961 GTGCAGAAAGCCCAACCCGACTGCAAGAGAGATCTGCGGCGCTGCGGCGGCGGAGAG 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysTrpIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGGAGAGAGATGATGACCGCTTGCAGAGGCTGGAGCGGCGGCGGCAAGAGCGCGG 1080

```

```

Db 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyLysProGluLysLeuValArgVal 359
QY 1081 CTGGCGGAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
Db 360 LeuAlaGluAlaMetSerGlnAlaAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe 379
QY 1138 AAGGCGCGCGGCGGATGTCGAAGTCTTCAACTGCGGCGGAGAGAGAGAGAGAGAGAG 1197
Db 380 LysGlyProArgArgIleValLysCysPheAsnGlyGlyGluGluGlnHisIleAlaArg 399
QY 1198 AACTCGCGCGCGCGCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
Db 400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGluGlnHisMet 419
QY 1258 AAGAGTGCACCGGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
Db 420 LysAspCysTrpGluArgGlnAlaAsnMetLeuGlyLysIleTrpProSerHisLysGly 439
QY 1318 CGCCCGGCAACTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
Db 440 ArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 CGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 460 ArgPheGluGluThrThrProAlaProLysGlnGluProIleGluArgGluProLeuThr 479
QY 1438 AGCCTGAAGAGCGTGTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1496
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492

```

#### RESULT 6

US-10-494-131-36  
Sequence 36, Application US/10494131  
Publication No. US20050176929A1

#### GENERAL INFORMATION:

```

/ APPLICANT: Williamson, Carolyn
/ APPLICANT: Van Harmelen, Joanne Heidi
/ APPLICANT: Gray, Clive Maurice
/ APPLICANT: Bourn, William
/ APPLICANT: Karim, Salim Abdool
/ TITLE OF INVENTION: HIV-1 Subtype Isolates Regulatory/Accesory Genes, and
/ FILE OF INVENTION: Modifications and Derivatives Thereof
/ FILE REFERENCE: 45669-300571
/ CURRENT APPLICATION NUMBER: US/10/494,131
/ CURRENT FILING DATE: 2004-04-30
/ PRIOR APPLICATION NUMBER: PCT/IB02/04550
/ PRIOR FILING DATE: 2002-10-31
/ PRIOR APPLICATION NUMBER: ZA 2001/8978
/ PRIOR FILING DATE: 2001-10-31
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 36
/ LENGTH: 492
/ TYPE: PRT
/ ORGANISM: Human immunodeficiency virus type 1
/ US-10-494-131-36

```

#### Alignment Scores:

```

Pred. No.: 6,61e-121 Length: 492
Score: 2448.00 Matches: 461
Percent Similarity: 97.36% Conservative: 19
Best Local Similarity: 93.51% Mismatches: 11
Query Match: 88.31% Indels: 2
DB: 18 Gaps: 2

```

US-09-475-704a-3 (1-1479) x US-10-494-131-36 (1-492)

```

QY 1 ATGGGCGCGGGCGGAGCATCTGGCGGGGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspLysTrpGlyLysIleArg 20
QY 61 CTGGCGCGCGGCGGCAAGAGTGTCTACATGATGAAGCACTGTGTGGGCGAGCGCGGAG 120

```

```
Db      21 LeuArgProGlyGlyLysLysHisIeTyMetLeuLysHisIeValTrpAlaSerArgGlu 40
Qy      121 CTGGAGAAATTGGCCCTGAAACCCCGGCTGCTGGAGACCAAGGAGGGCTGCACAAGCATC 180
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGlnGlyCysLysGlnIle 60
Qy      181 ATCCGCGACGTCACACCCCGGCTGCAGACCGGAGGAGGAGAGCTGTAAGAGCCGTTCACAC 240
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 MetLysGlnLeuGlnProAlaLeuGlnmTrGlyThrGlnGluLeuLysSerLeuTyraen 80
Qy      241 ACCGTGGCCACCCCTTACTGCTGTGACAGAAAGATGAGGTCCGCGACACCAAGAGGCC 300
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      81 ThrValAlaThrLeuTyrcysValHisGlnLysIleGluValArgAspThrLysGlnAla 100
Qy      301 CTGGACACAGATGAGAGAGAGCAAGAACATGTCACAGACAAAGATCCAGAGGCCCGAGGCC 360
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      101 LeuAspLysIleGlnGlnGlnGlnGlnmLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
Qy      361 GCCGACAAAGGAGGAGGTGAGCCAGAACTACCCCATGTGTCAGAAACCTGACAGGCGCCAGATG 420
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 AlaAsp--GlyLysValSerGlnAsnTyrrProIleValGlnAsnLeuGlnGlyGlnMet 139
Qy      421 GTGACACAGGACCATGAGCCCGGCAACCTGAAAGCCTGGGTGTAAGGTGATCGAGAGAAAG 480
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGlnGlnLys 159
Qy      481 GCGTTTCAGCCCGGAGGTGATCCCATGTTTACCCGCTGAGCGAGGAGGCGCCACCCCGCAG 540
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      160 AlaPheSerProGlyValIleProMetcPheThrAlaLeuSerGlnGlyAlaThrProGln 179
Qy      541 GACCTGAACAGATGTTGAACAACCGTGGGCGGCGCCACAGCGGCGCATGACATGCTGAAG 600
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      180 AspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaIleMetcIleMetLeuLys 199
Qy      601 GACACCATCAACGAGAGGCGCGGAGTGGAGCCGCTGACACCCGCTGACACCCGCGGCGCC 660
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      200 AspThrIleAsnGlnGlnAlaIleGlnTrpAspArgValHisProValHisIleGlyPro 219
Qy      661 ATGCGCCCGGCGCAATGCGGAGCGCCCGGCGAGCGACATGCGCGGACACCAACGACCC 720
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      220 IleAlaProGlyGlnMetArgGlnLupProArgGlySerAspIleAlaGlyThrThrSerThr 239
Qy      721 CTGGAGAGAGATGCGCTGATGACCGACGACCCCGGCTGCGGCGGCGACATCTAC 780
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      240 LeuGlnGlnGlnIleAlaTrpMetcThrSerAsnProProIleProValGlyAspIleTyrr 259
Qy      781 AAGCGGTGATCATCTGGGCTGTAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 840
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      260 LysAspGTrpIleIleLeuGlyLeuAsnLysIleValAlaGlnMetcLysSerProValSerIle 279
Qy      841 CTGACATCAACGAGGCGCGGCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 900
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      280 LeuAspIleArgGlnGlyProLysGlnLupProPheArgAspTyrrValAspArgPheLys 299
Qy      901 ACCCTGGCGCGGCGAGAGGACCAAGAGGTGAAGAACTGTATACCGACACCTTCTGCTG 960
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      300 ThrLeuArgAlaGlnGlnAlaThrGlnGlnValLysAsnTrpMetcThrAspThrLeuLeu 319
Qy      961 GTGCAAGACGCGCAACCCGCTGCAAGACCATCTGCGGCTCTGCGGCGCGGCGCGCAGC 1020
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
Qy      1021 CTGAGAGAGATGATACCGCTGTCAGAGGCGGTGGCGCGGCGCCAGCCACAGGCGCGCTG 1080
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      340 LeuGlnGlnMetcMetcThrAlaCysGlnGlnValGlyGlyProGlyHisIleValAlaArgVal 359
Qy      1081 CTGGCGCGAGCGGATGAGCCAGGCGCAACAC--AACGTGATGATGACAGAGCAACTTC 1137
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      360 LeuAlaGlnAlaMetSerGlnAlaAsnSerGlyAlaAsnIleMetcGlnArgSerAsnPhe 379
Qy      1138 AAGGCGCGCGGCGGATGCTCAAGTGTTCATCTGCGGAGAGAGGAGGCGCACTCCGCGCG 1197
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      380 LysGlyProArgArgIleValLysCysPheAsnCysGlyLysGlnGlyHisIleIleAlaArg 399
```

```
Qy      1198 AACTCCCGCGCGCGCGCGGCGAAGAGGCTGTGAAAGTCCGCGCAAGAGGCGCCACCATG 1257
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGlyHisIleGlnMet 419
Qy      1258 AAGGATCAACCGAGCGCGGCGCAACTTCCGTGGCGCAAGATCTGGCGCGGCGCAAGGCGC 1317
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      420 LysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 439
Qy      1318 CGCGCGGCGCACTTCTCTGACAGCGCGCGGCGGCGGCGCGCGCGCGCGCGGAGAGCTTC 1377
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      440 ArgProGlyAsnPheLeuGlnAlaAsnArgProGlnProThrAlaProProAlaGlnSerPhe 459
Qy      1378 CGCTTCGAGAGACCAACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1437
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      460 ArgPheGlnGlnGlnThrThrProAlaProLysGlnGlnLupProIleGluArgGluProLeuThr 479
Qy      1438 AGCCTGAAGACCTGTTGCGCAAGAGCCCGCTGAGCGCG 1476
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      480 SerLeuLysSerLeuPheGlySerAspProLysSerGln 492

RESULT 7
US-09-991-258-5
; Sequence 5, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Ketch, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 492
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-5

Alignment Scores:
Pred. No.: 3,19e-120 Length: 492
Score: 2435.00 Matches: 458
Percent Similarity: 96.96% Conserved: 20
Best Local Similarity: 92.90% Mismatches: 13
Query Match: 87.84% Indels: 2
Gaps: 2

US-09-475-704A-3 (1-1479) x US-09-991-258-5 (1-492)
Qy      1 ATGGGCGCGCGGCGAGCATCTGCGGCGGCGGCAAGTGAAGCGCTGGAGGCGATCCGC 60
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MetAlaAlaArgAlaSerIleLeuArgGlyGlnLysLeuAspLysTrpGlnLysIleArg 20
Qy      61 CTGGCGCGCGGCGGCAAGAACTGTACATGATGAAGCACTGTGTGTGGCGGCGGCGAG 120
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      21 LeuArgProGlyGlyLysLysHisIeTyMetLeuLysHisIeValTrpAlaSerArgGlu 40
Qy      121 CTGGAGAAATTGGCCCTGAAACCCCGGCTGCTGGAGACCAAGGAGGGCTGCACAAGCATC 180
```



Db	41	LeuGIuAIGPhReAlaLeuAsnProGIyLeuLeuGIuIntrSerGIuGIyCyLySeGIuIle	60
Oy	181	ATCCGCCAGCTGCACCCGCCCTTCGCAGACCGGACCGAGAGCTGAAGAGCTTGTCAC	240
Db	61	MetLysGIuLeuGIuInProAlaLeuGIuInThrGIyThrGIuGIuLeuLySerLeuTyAsn	80
Oy	241	ACCGGGGCCACCCCTGACTGCTGTCAGACGAAGAATCCGAGTCCCGGACACCAAGAGGCC	300
Db	81	ThrValAlaTrpLeuTyCyValAlHISGIuLySIIeGIuValAArgAspThrLySeGIuAla	100
Oy	301	CTGCACAAAGATCGAGAGAGACAGAACAAAGTGCAGACGAAGATCTCAGACCGGACGCC	360
Db	101	LeuAspLySIIeGIuGIuGIuGIuInAsnLySeCySeGIuInGIuLyThrGIuInAlaLySIIa	120
Oy	361	GCCGACAAAGGGCAGGTAGCCAGAACCTAACCCCATCTGCAGAACCTTCAGGGCCAGATG	420
Db	121	AlaAsp---GIyLyValSerGIuInAsnTyProIIeValAlGIuInAsnLeuGIuInGIuMet	139
Oy	421	GTGCACCAAGGCATCAGGCCCGGCCGACCTTGAAACGCTGGGTGAAGGTATGAGAGAGAG	480
Db	140	ValHISGIuAlaIIeSerProArgThrLeuAsnAlaTrpValLySIIaIIeGIuInLyS	159
Oy	481	GCTTTCAGCCCCGAGGTATCCCATGTTCACCGCCCTTGACCGAGGGGCCACCCGCCAG	540
Db	160	AlaPheSerProGIuValIIeProMetPheThrAlaLeuSerGIuGIyAlaTrpProIIIn	179
Oy	541	GACCTGAACACGATGTTGAACACGCTGGGGGCCGACCGAGGCCGACATGCTGAG	600
Db	180	AspLeuAsnTrpMetLeuAsnThrValaIGyGIyHISGIuAlaAlaMetGIuMetLeuLyS	199
Oy	601	GACACCATCAACGAGAGAGCGGCCGAGTGGGACCGCGCTGCACCCCGTGCACGGCGGCC	660
Db	200	AspThrIIeAsnGIuGIuAlaAlaGIuTrpAspArgLeuHISProValHISaIIeGIyPro	219
Oy	661	ATCGCCCCCGGCGCAATGCGGAGGCCCGCGCGGACGACATCGCGCGGACCAACGACGC	720
Db	220	IIeAlaProGIyGIuMetArgGIuProArgIySerAspIIeAlaGIyThrThrSerThr	239
Oy	721	CTGCAGAGCAGATGCGCTGTGATGACGACAAACCCCATCCCGCTGGGAGCATCTAC	780
Db	240	LeuGIuInGIuInIIeAlaITrpMetThrSerAsnProProIIeProValaGIyAspIIeTyx	259
Oy	781	AAGCGGTGATCATCTTGAGGCGCTGAACAGATCGTGGAGTGAAGCCCGCTGAGCATC	840
Db	260	LySAspTrpIIeIIeLeuGIyLeuAsnLySIIeValAArgMetTySerProValSerIIe	279
Oy	841	CTGCACATCAACGAGAGGCCCAAGAGGCGCTTCGCGGACTGAGGACCGGCTTCTTCAG	900
Db	280	LeuAspIIeAspGIuInGIyProLySIIeProPheArgAspTyValaAspArgPhePheLyS	299
Oy	901	ACCGTCCGCGCGAGCAGAGACGCCAGAGAGGTGAAGAACTGAGATCCGAGACCCCTGCTG	960
Db	300	ThrIIeAspArgIIaGIuGIuAlaThrGIuInGIuValLySAsnTrpMetThrAspThrLeuLeu	319
Oy	961	GTGCAGAACGCCAACCCCGACTGCAAAACCATCTTCGCGGCTCTTGCGCCCCGGCGGACG	1020
Db	320	ValGIuInAsnAlaAsnProAspCyValThrIIeLeuAspArgIIaLeuGIyProGIyAlaThr	339
Oy	1021	CTGCAGAGATGATATCAGCGCTGCGCAGAGGGCGTGGCGGCCGACGAAAGGCGCGGTG	1080
Db	340	LeuGIuGIuMetMetThrIIaCySeInaIGyAlaIGyLyProGIuInHISLyValaArgVala	359
Oy	1081	CTGGCCGAGGCGATGAGCCGAGGCCAACAC---AGCGTGAATGATCAGAGAGCAATCTC	1137
Db	360	LeuAlaGIuAlaMetSerGIuInThrAsnSerGIyAsnIIeMetMetGIuInAspSerAsnPhe	379
Oy	1138	AAGGAGCCCCCGGCGGATGTCAGAGGCTTCAACTGCGGACAGAGAGGCGACATGCGCGCG	1197
Db	380	LySeGIyProArgArgIIeValaLySeCyPheAsnCySeGIyLySIIeGIuInHISIIeAlaArg	399
Oy	1198	AACCTGCGCGCCCCCGGACGAAGGAGCTCTGGAAGTGCAGGACGAGGCGCACCATGATG	1257

```

Db      400  AaenCySaAgAlaPrOaRgLyLvgSgLyCySrTrPrLyScySgLyLvgLgUgLyHIselInetC 419
Qy      1258  AAGAGCTGCACCGAGCGCCAGGCCAATTCTCTGGGCAAGATCTGGCCCAAGCCCAAGAGGC 1317
Db      420  LysAspCysThrIguArgGlnAlaAsnPhenLeuGlyLysIleTrpProSerHisIleSgLy 439
Qy      1318  CGCCCCGGGCAACTTCTCTGAGAGCGCGCCGAGGCCACCGGCCCCCGCCGAGAGCTTC 1377
Db      440  ArgProGlyAsnPhenLeuGlnAsnMyrProGluProIrrAlaProPalaGluSerPhe 459
Qy      1378  CGCTTCGAGAGACCAACCCCGCCGAGCAAGAGAGAGCAAGACCGGAGACCTGACC 1437
Db      460  ArgPheGluGluThrThrProAlaProLysGlnGluProIleGluArgLupProLeuThr 479
Qy      1438  AGCCGGAAGACCTGTTGGGCAAGACCAACCCCTGAGCCGAG 1476
Db      480  SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492

RESULT 8
US-10-929-234-5
; Sequence 5, Application US/10929234
; Publication No. US20050123555A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Caley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanslstrom, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 0113, 0001u3
CURRENT APPLICATION NUMBER: US/10/929,234
CURRENT FILING DATE: 2004-08-30
PRIOR APPLICATION NUMBER: US/09/991,258
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 492
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence, Note =
US-10-929-234-5

Alignment Scores:
Pred. No.:      3,19e-120      Length:      492
Score:          2435.00      Matches:      458
Percent Similarity: 96.96%      Conservative: 20
Best Local Similarity: 92.90%      Mismatches:  13
Query Match:    87.84%      Indels:      2
DB:            18      Gaps:      2

US-09-475-704A-3 (1-1479) x US-10-929-234-5 (1-492)
Qy      1  ATGGAGCGCCGCGCCAGCATCTCTGCGCGCGGCAAGCTGAGCGCTTGGAGCGCATCCGC 60
Db      1  MetaLaalaaAgAlaSerIleLeuArgSgLyGluLysLeuAspLysTrpGluLysIleArg 20
Qy      61  CTGAGCGCCGCGGCGGCAAGAGGCTACATGATGAAGACCATGCTGTGGAGCCGAGCGCGAG 120
Db      21  LeuArgProGlyGlyLysLysHisIleTrpMetLeuLysHisIleValATrpAlaSerArgGlu 40
Qy      121  CTGAGAGAGTTTCGCCCTGAACCCCGGCTGCTGTGAGAGACAGGAGGCTGCAGAGAGATC 180

```





QY	241	ACCGTGGCCACCCCTTACTGCGGTGACAGAAAGATCGAGGTCCTGGACACCAAGAGGCC	300
Db	111	ThrValAlaTrpTrpLeuTrpCysValHisGluTrpHisLeuValAlaArgAspThrIleGluAla	130
QY	301	CTGGCAAGATCTCGAGGAGAGAGACAGAAACAAGTCCAGCAGAGATCCACAGGCCAGAGCC	360
Db	131	LeuAspArgHisIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	150
QY	361	GCCGCAAGGGCAGGTGAGGCCAGACTACCCCATCTGTGCAGAACCTTGAGGGGCCAGATG	420
Db	151	AlaAsp---GlyIleValSerGlnAsnTrpProIleValGlnAsnLeuGlnGlyIleMet	169
QY	421	GTGGCAAGGCGCATGAGCCCGCCAGCCCTGAAACGGCTGGGGTGAAGGTATCGAGAGAG	480
Db	170	ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIleValIleGluGluTrp	189
QY	481	GCCTTGACGCCCGAGGTATCCCATGTTCACCGGCTTGAGCGAGGGCGCCACCCCGCAG	540
Db	190	AlaPheSerProGluValIleIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln	209
QY	541	GACCTGAACACGATGTTGAACAACCGTGGCGGGCCACAGGCCCATGCAATGCTGAAG	600
Db	210	AspLeuAsnTrpMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuTrp	229
QY	601	GACACCATCAACGAGAGAGGCGCGCAGTGGGAGACCCCGTGCACCCCGTCACAGCCGAGCCC	660
Db	230	AspThrTrpIleAsnGluGluAlaIleGluTrpAspArgLeuHisIleProValHisIleGlyPro	249
QY	661	ATCGCCCCCGGCAGATGGCGAGGCCCGCGCGGACGACATGCGCGGACACCAACGACAC	720
Db	250	IleAlaProGlyGlnMetArgIleProArgIleSerAspIleAlaGlyThrThrSerThr	269
QY	721	CTGGCAAGACGATGCGCTGGATGACGACGAAACCCCGCATCCCGTGGAGCAATCTAC	780
Db	270	LeuGlnGluGlnIleAlaIleTrpMetThrSerAsnProPoleProValGlyAspIleIleTrp	289
QY	781	AAGCGGTGATCATCTCGGGCCCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC	840
Db	290	LysArgTrpIleIleLeuGlyLeuAsnLysIleValAlaGlyMetTyrSerProValSerIle	309
QY	841	CTGGCAATCAAGCAGAGGGCCCAAGAGACCTTCGCGATACGTGACCGGCTTTCTACG	900
Db	310	LeuAspIleArgGlnGlyProLysGluProPheArgAspTrpValAspArgPheTrpHis	329
QY	901	ACCCGCGCGCGCAGACGACGACCAACGAGAGGTGAAGAACTGGATGACCGACACCCCTGCTG	960
Db	330	ThrIleuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu	349
QY	961	GTGCAGAACCGCAACCCCGCATCTGCAAGAACATCTTGGCGGCTTGGCGCCGCGCCAGC	1020
Db	350	ValGlnAsnAlaAsnProAspCysIleThrIleLeuArgAlaLeuGlyProGlyAlaThr	369
QY	1021	CTGGAGAGATGATGATACCGCTGCGCAGAGGGCGGGCGGCGCCAGCAACAAGCCCGCGTG	1080
Db	370	LeuGluGluMetMetThrIleCysGlnGlyValGlyIleProGlyHisIleValAlaArgAla	389
QY	1081	CTGGCCGAGCGGATGAGCCAGGCCAACAC---AGCGTATGATCAAGAGCAATCTC	1133
Db	390	LeuAlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe	409
QY	1138	AAGGGCCCCCGCGGATGTCACAGGCTTCAACTCGCGCAAGAGAGGCGACATGCGCCGC	1199
Db	410	LysGlyProArgArgIleValIleCysPheAsnCysGlyLysGluGlyHisIleAlaArg	429
QY	1198	AACTGCGCGCGCCCCCGCAGAAAGGGCTGCTGGAAGTGCGGCAAGAGGGCCACCAAGATG	1257
Db	430	AsnCysValArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMet	449
QY	1258	AAGACATGACCGGACGCGCAGGCCAACTTCTGGGGCAAGATCTGGCCACGACCAAGAGGC	1317
Db	450	LysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisIleGlyGly	469

Oy	1318	GGCCCCGGCACTTCCTGACAGCGCCGCCAGCGCCACCGCCCCCGCGGAGAGCTTC	1377
Db	470	ArgProGlyaAnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe	489
Oy	1378	CGCTTCGAGAGACCAACCCCGCGCGACAGACGAGACGACCGGACCGGAGACCTTGACC	1437
Db	490	ArgPheGluGluThrThrProAlaProIleArgIleGluArgIleGluProLeuThr	509
Oy	1438	AGCCTGAAGACCTGTTGGCGACGACCCCTGAGCCAG	1476
Db	510	SeuLeuYSerLeuPheGlySerAspProLeuSerGln	522
RESULT 10			
US-10-339-217-110			
Sequence 110, Application US/10339217			
Publication No. US20030198648a1			
GENERAL INFORMATION:			
APPLICANT: Buechter, Douglas			
APPLICANT: Hou, Xiaohong			
APPLICANT: Marlor, Christopher W.			
APPLICANT: Rice, William G.			
APPLICANT: Yang, Wengang			
TITLE OF INVENTION: Methods for Identifying Compounds which Inhibit Binding of			
TITLE OF INVENTION: Nucleospsid 7 Protein to HIV-1 RNA			
FILE REFERENCE: 111021.143(ACH-US1)			
CURRENT APPLICATION NUMBER: US/10/339,217			
PRIOR FILING DATE: 2003-01-09			
PRIOR APPLICATION NUMBER: US 60/347,369			
PRIOR FILING DATE: 2002-01-11			
NUMBER OF SEQ IDS: 144			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 110			
LENGTH: 492			
TYPE: PRT			
ORGANISM: HIV-1			
US-10-339-217-110			
Alignment Scores:			
Pred. No.:	1.37e-119	Length:	492
Score:	2423.00	Matches:	451
Percent Similarity:	96.96%	Conservative:	27
Best Local Similarity:	91.48%	Mismatches:	13
Query Match:	87.41%	Indels:	2
DB:	14	Gaps:	2
US-09-475-704A-3 (1-1479) x US-10-339-217-110 (1-492)			
Oy	1	ATGGGCGCGCGCGCGACGATCTGTGGCGCGCGGACGCTGGAAGCGCTGGAGCGCATCCGC	60
Db	1	MetGlyAlaArgLAserLLeuArgGlyGluLysLeuAspThrTrpGluLysIleArg	20
Oy	61	CTGGCGCGCGCGGCGGACAGAGTGTCAATGATGAAGACACTGTGTGGCGCGACCGCGAG	120
Db	21	LeuArgProGlyGlyLysLysCysTyrMetLeuLysHisIleValTrpAlaSerArgL	40
Oy	121	CTGGAAGAATTGCGCCCTGAACCCCGCGCTGTGAGACCGACGAGAGGCTGCAAGCATC	180
Db	41	LeuGluArgPheSerLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle	60
Oy	181	ATCGCGCAGCTGCACCCCGCGCTGCGAGACCGGACGAGAGAGCTGAAGAGCTGTTCAC	240
Db	61	MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn	80
Oy	241	ACCGTGGCCACCTGTACTGTGCGGTGACGAGAAATCGAGTCCGGGACCAAGAGAGGCC	300
Db	81	ThrValAlaThrLeuPheCysValHisGluLysIleAlaValArgAspThrLysGluAla	100
Oy	301	CTGCAACAAGATTCAGAGAGAGAGCAACAAGATCCACAGAAATCCAGCAGGCGCGAGCC	360
Db	101	LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnIleThrGlnIleAlaLysAla	120
Oy	361	GCCGCAAGGGCGCAAGGTAGACCAAGAACTAACCCCATGTGACAGAACTGTGAGGGCCAGATG	420

Db 121 Alaasp---GlyThrValSerGlnAsnTyrrProIleValGlnAsnLeuGlnGlnMet 139  
QY 421 GTGCACCAAGGCGCATCAGCCCGCGACCTGTGAACGCTGTGGTGAAAGTGATCGAAGGAAG 480  
DB 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIleGlnIuIuys 159  
QY 481 GCCTTCAAGCCCGAGGTGATCCCATGTTCAACCGGCTGTGACAGAGGCGCCACCCCGCAG 540  
DB 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGlnIuIuValaThrProGln 179  
QY 541 GACCTGAACAGATGTTGAACAACCTGTGGCGGCGCACAGGCGCCACGTGAGATGTCTGAAG 600  
DB 180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuIuys 199  
QY 601 GACACCATCAAGAGAGGCGCCGCGAGTGGAGACCGGTGTGACCCCGGTGACCGCGGCGCC 660  
DB 200 AspThrIleAsnGlnIuIuValaIaGlnTrpAspArgLeuHisIleProValHisIaGlyPro 219  
QY 661 ATCGCCCGCGGCGAGATGGCGGCGCGCCGCGGCGAGATGGCGGCGCCACCAAGCAGC 720  
DB 220 IleAlaProGlyGlnMetArgIuProArgIuSerAspIleAlaGlyThrThrSerThr 239  
QY 721 CTGCAGAGACAGATCGCTGTGATGACCAACACCCCGCATCCCGTGGCGACATCTTAC 780  
DB 240 LeuGlnGlnGlnIleAlaTrpMetThrAsnAsnProProValProValGlyAspIleTyrr 259  
QY 781 AAGCGGTGATCATCTGTGGGCTGTGAACAAGATGTGCGGATGTACAGCCCGTGAACATC 840  
DB 260 LysArgTrpIleIleLeuGlyLeuAsnLysIleValaIArgMetCysSerProValSerIle 279  
QY 841 CTGGACATCAAGAGGCGCCCAAGAGCGCTTCCGCGACATGACGTGACCGCTTCTTCAG 900  
DB 280 LeuAspIleArgGlnIuIuProLysGlnProPheArgAspIuValaAspArgPheLeuLys 299  
QY 901 ACCCTGCGGCGCGAGAGAGACCCAGAGGTGAAGAACTGATGACCAACCCGCTG 960  
DB 300 ThrIleArgAlaGlnGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuAsn 319  
QY 961 GTGCAGAACGCCAACCCCGCATGTGCAAGACATCTTGCGGCTTGTGGCGCGCGCAGC 1020  
DB 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339  
QY 1021 CTGGAGAGATGATACCCGCTGCGAGGCGGTGGCGGCGCCAGCCACCAAGGCGCGCTG 1080  
DB 340 LeuGlnGlnMetMetThrAlaCysGlnGlyValGlyLysProSerHisLysAlaArgVal 359  
QY 1081 CTGGCGAGGCGATGAGCGAGCGC---AACAACAAGGTGATGATGACGAAGCAACTTC 1137  
DB 360 LeuAlaGlnAlaMetSerGlnThrAsnAsnAlaAsnIleMetGlnArgSerAsnPhe 379  
QY 1138 AAGGCGCCCGCGGCGATCTGTCAAGTCTTCAACTGTGCGGAGAGAGGCGCATCGCCGCG 1197  
DB 380 LysGlyProArgArgIleIleLysCysPheAsnGlyLysGlnIuIuHisLeuAlaArg 399  
QY 1198 AACTCGCGGCGCGCCCGCAAGAGGCTCTGTGAAGTGGCGCAAGAGGCGCCACCAAGT 1257  
DB 400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnIuIuHisGlnMet 419  
QY 1258 AAGGACTGACAGGCGCGCAACTTCTGTGGGCAAGATCTGGCGCGCCACCAAGGCGC 1317  
DB 420 LysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 439  
QY 1318 GCCTCCGCGCACTTCTGTGACGCGCGCCGAGCCGCGCCCGCGCGAGAGCTTC 1377  
DB 440 ArgProGlyAsnPheLeuGlnAsnArgProGlnProThrAlaProProAlaGlnSerPhe 459  
QY 1378 CGCTTGAAGAGACCAACCCCGCGCAAGAGCAAGAGCAAGAGACCGCGACCTGTAGC 1437  
DB 460 ArgPheGlnGlnThrThrProAlaProLysGlnGlnProArgIuGlnProLeuThr 479  
QY 1438 AGCTTAAGAGCCTGTTGGCAACGACCCCGTGAAGCAG 1476  
DB 480 SerIleuLysSerIleuPheGlySerAspProLeuSerGln 492

RESULT 11  
US-10-339-217-109  
; Sequence 109, Application US/10339217  
; Publication No. US20030198648A1  
; GENERAL INFORMATION:  
; APPLICANT: Buechter, Douglas  
; APPLICANT: Hou, Xiaohong  
; APPLICANT: Marlor, Christopher W.  
; APPLICANT: Rice, William G.  
; APPLICANT: Yang, Wengang  
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of  
; FILE REFERENCE: 111021.145 (ACH-US1)  
; CURRENT APPLICATION NUMBER: US/10/339,217  
; PRIOR FILING DATE: 2003-01-09  
; PRIOR APPLICATION NUMBER: US 60/347,369  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 109  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: HIV-1  
US-10-339-217-109  
  
Alignment Scores:  
Pred. No.: 3,19e-119 Length: 492  
Score: 2416.00 Matches: 455  
Percent Similarity: 95.74% Conservative: 17  
Best Local Similarity: 92.29% Mismatches: 19  
Query Match: 87.16% Indels: 2  
DB: 14 Gaps: 2  
  
US-09-475-704A-3 (1-1479) x US-10-339-217-109 (1-492)  
QY 1 ATGGGCGCGCGCGCGAGCATCTGTGCGCGCGCGCAAGCTGTGACGCTGTGGAGCGCATCCGC 60  
DB 1 MetGlyAlaSerAlaSerIleLeuArgIlyGlyLysLeuAspLysTrpGlnLysIleArg 20  
QY 61 CTGGCGCGCGCGCGCAAGAGTGTACATGTGAAGCACTGTGTGGCGCGCGCGAG 120  
DB 21 LeuArgProGlyGlyLysLysLysTyrrArgLeuLysHisLeuValTrpAlaSerArgGln 40  
QY 121 CTGGAAGATTGCGCTTGAACCCCGCGCTGTGAGACCAAGAGCGCTGCAAGCAGATC 180  
DB 41 LeuGlnArgPheAlaLeuAsnSerGlyLeuLeuGlnThrAlaGlnGlyCysLeuGlnIle 60  
QY 181 ATCGCGAGCTGTGACCCCGCGCTGTGACACCGGACCGAGAGCTGAAGCTGTTCAC 240  
DB 61 IleLysGlnLeuGlnProAlaLeuGlnThrGlyThrGlnGlnLeuLysSerIleuTyrrAsn 80  
QY 241 ACCGCGGCGACCGTGTACTGTGGTGCAGAGAAGATCGAGTCCGCGACCAAGAGAGCGC 300  
DB 81 ThrValAlaThrLeuTyrrCysValHisAlaGlyIleGlnLysAlaArgAspThrLysGlnAla 100  
QY 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTGTCCAGCAAGAAGATCCACAGCGCGAGCC 360  
DB 101 LeuAspLysIleGlnGlnGlnGlnAsnLysCysGlnGlnLysTrpGlnGlnAlaLysGln 120  
QY 361 GCGGCAAGGCGAGAGTGAAGCAAGACTTACCCCATGTGTGCAAGCTGTGAGGCGCGCATG 420  
DB 121 Alaasp---GlyLysValSerGlnAsnTyrrProIleValGlnAsnLeuGlnGlyGlnMet 139  
QY 421 GTGCACAGGCGCATAGCGCCCGGACCTGTGAAGCTGTGGTGAAGGTATGAGAGAGAG 480  
DB 140 ValHisGlnProIleSerProArgThrLeuAsnAlaTrpValIleGlnIuIuys 159  
QY 481 GCCTTCAAGCCCGAGGTATCCCATGTTCAACCGGCTGTGACAGAGGCGCCACCCCGCAG 540  
DB 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGlnIuIuValaThrProGln 179  
QY 541 GACCTGAACAGATGTTGAACAACCGTGGCGGCGCACAGGCGCCACGTGAGATGTCTGAAG 600





```

QY 901 ACCCTGCGCGCCGAGCAGACACCCAGAGTGAAGAACTGATGACCCGACACCTGTCTG 960
    |||
Db 300 ValLeuAArgAlaGluInAlaThrGlnAaPValLysAntzPMeThrAphrLeu 319
    |||
QY 961 GTGCAGAACGCCCAACCCCGCATCTGCAAGACATCTGCGCTTCGGCCCGCGCCGACG 1020
    |||
Db 320 ILeGlnAaPValLysAntzPMeThrAphrLeu 339
    |||
QY 1021 CTGGAGGAGATGATGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
    |||
Db 340 LeuGlnGlnMetMetThrAlaCysGlnGlnValGlyGlyProGlyHsAlaVal 359
    |||
QY 1081 CTGGCGGAGCGATGAGCAGCGCCCAACACACAGCGTATGATGACAGAAAGCAATTCAG 1140
    |||
Db 360 LeuAlaGlnAlaMetSerGlnAlaAaSerAaLleMetCecDlnAArgSerAaPheLys 379
    |||
QY 1141 GGCCTCCCGCGCATCTGCAAGTCTTCAACTGCGCGCAAGAGCGCGCATCTGCGCGCAAC 1200
    |||
Db 380 GlySerLysArgIleValLysCysPheAaCysGlyGlnGlyLysGlyLysIleAlaArgAa 399
    |||
QY 1201 TGCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
    |||
Db 400 CysAArgAlaProAArgLysLysGlyCysTrpLysCysGlyGlnGlnLysIleGlnMetLys 419
    |||
QY 1261 GACTGACACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
    |||
Db 420 AaPArgTrpGlnArgGlnAlaAaPheLysGlyLysIleTrpProSerHsAlaVal 439
    |||
QY 1321 CCGCGCACTTCTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
    |||
Db 440 ProGlyAaPheLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 459
    |||
QY 1381 TTGCGAGGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
    |||
Db 460 PheGlnGlnTrpThrProAlaProLysGlnGlnLysLysAaPArgLysProLeuLysLys 479
    |||
QY 1441 CTGAAGAGCGCTGTTGCGCAACGACCCCTGAGCGCAG 1476
    |||
Db 480 LeuLysSerLeuPheGlySerAaPProSerSerGln 491
    |||

RESULT 14
US-10-339-217-143
; Sequence 143, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buecher, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Method for Identifying Compounds Which Inhibit Binding of
; FILE REFERENCE: Nucleocapsid 7 Protein to HIV-1 RNA
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143
; LENGTH: 491
; TYPE: PR
; ORGANISM: HIV-1
US-10-339-217-143

Alignment Scores:
Pred. No.: 3,39e-118 Length: 491
Score: 2396.50 Matches: 449
Percent Similarity: 95.53% Conservative: 21
Best Local Similarity: 91.26% Mismatches: 21
Query Match: 86.45% Indels: 1
DB: 14 Gaps: 1

```

```

US-09-475-704A-3 (1-1479) x US-10-339-217-143 (1-491)
QY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
    |||
Db 1 MetGlyAlaArgAlaSerLysLeuAArgGlyGlyLysLeuAaPValTrpGlnLysLysArg 20
    |||
QY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    |||
Db 21 LeuAArgProGlyGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 40
    |||
QY 121 CTGGAGAAATTCGCTGAAACCCCGCGCTGCGAGACACCGAGCGCGCGCGCGCGCGCG 180
    |||
Db 41 LeuGlnAArgPheAlaLeuAaPProGlyLeuLeuGlnTrpSerGlnGlyCysLysGlnLys 60
    |||
QY 181 ATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
    |||
Db 61 IleLysGlnLeuGlnProAlaLeuGlnTrpGlnGlnLysLysLysLysLysLysLysLys 80
    |||
QY 241 ACCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
    |||
Db 81 ThrValAlaTrpLeuTrpCysValHisAlaGlyIleGlnLysLysLysLysLysLysLys 100
    |||
QY 301 CTGACCAAGATTCGAGAGAGACAGAAAGTGCACAGAAAGATTCACAGCGCGCGCGCG 360
    |||
Db 101 LeuAaPArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 120
    |||
QY 361 GCGCGCAAGGCGCGAGTGAAGCAACCACTGCGCGCAAGCTGCGCGCGCGCGCGCGCG 420
    |||
Db 121 AlaAaP--GlyLysValSerGlnAaPProLysLysLysLysLysLysLysLysLysLys 139
    |||
QY 421 GTGCACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
    |||
Db 140 ValHisGlnAlaLeuSerProAArgThrLeuAaPValLysValLysGlnLysLys 159
    |||
QY 481 GCTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
    |||
Db 160 AlaPheSerProGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 179
    |||
QY 541 GACCTGAACAGAGTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
    |||
Db 180 AaPLeuAaPTrpMetLeuAaPThrValGlyGlyHisGlnAlaAlaMetGlnMetLysLys 199
    |||
QY 601 GACACCATCAACGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
    |||
Db 200 AaPTrpTrpLeuGlnGlnAlaGlnLysTrpAaPArgLysLysLysLysLysLysLysLys 219
    |||
QY 661 ATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
    |||
Db 220 IleAlaProGlyGlnMetAArgLysProAArgLysLysLysLysLysLysLysLysLys 239
    |||
QY 721 CTGCAGAGGAGATGCGCTGATGACAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 780
    |||
Db 240 LeuGlnGlnGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 259
    |||
QY 781 AAGCGGTGATCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
    |||
Db 260 LysAArgTrpLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 279
    |||
QY 841 CTGACATCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
    |||
Db 280 LeuAaPTrpLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 299
    |||
QY 901 ACCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
    |||
Db 300 ThrLeuAArgAlaGlnGlnAlaThrGlnAaPValLysAaPTrpMetThrAaPTrpLeuLys 319
    |||
QY 961 GTGCAGAACGCCCAACCCCGCATCTGCAAGACATCTGCGCTTCGGCCCGCGCCGACG 1020
    |||
Db 320 ValGlnAaPValLysAntzPMeThrAphrLeu 339
    |||
QY 1021 CTGGAGGAGATGATGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
    |||

```

Db 340 LeuGIuGIuMeTMeThrAlaCySGInGIyValGIyLProSeRHiSlYsAlaArGVal 359  
QY 1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCAAGCTGATGATGAGAAAGCAACTTCGAAG 1140  
Db 360 LeuAlaGIAlAlaMeSeRGIInThraSeRnASeRlIeUmeTGIInArGeRAnPheLys 379  
QY 1141 GGGCCCCGGGCGATGCTCAAGTCTTCACCTGGCGGAAGAGGGCCACATCCCGCCAGC 1200  
Db 380 GIyPheLysArGThrValIlyScYPheAnCySGIlyLysGIuGIyHlIeAlaArGAn 399  
QY 1201 TCCCGGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGGAAGAGGCGCCACCATGAAG 1260  
Db 400 CySaRgAlaProArGIySlYsGIyStRPlYsCySGIlyLysGIuGIyHlIeAlaArGAn 419  
QY 1261 GACTGACCGAGCGCGCAAGGCGCAACTTCCTGGGCAAGATCTGGCCGCAAGAGGCGCGC 1320  
Db 420 AspCyThrGIuArGIAlAlaPheLeuGIyLylIeTrProSeRHiSlYsGIyArG 439  
QY 1321 CCGCGCACTTCCTGAGAGCGCGCCCGGAGCCCGCCCGCCCGCGAGCTTCGCGC 1380  
Db 440 ProGIyAsnPheLeuInAsnArGProGIuProThAlaProProAlaGIuSeRPhArG 459  
QY 1381 TTCGAGAGACCAACCCCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Db 460 PheGIuGIuThThrThrProAlaLeuLysGIuGIuGIuLysArGIuProLeuThrSer 479  
QY 1441 CTGAAGAGCTGTTCGCAACGACCCCTGAGCCAG 1476  
Db 480 LeuLysSeRLeuPheGIySeRAsPProLeuSeRGIIn 491

## RESULT 15

US-10-339-217-108  
/ Sequence 108, Application US/10339217  
/ Publication No. US20030198648A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Buechter, Douglas  
/ APPLICANT: Hou, Xiaohong  
/ APPLICANT: Marlor, Christopher W.  
/ APPLICANT: Rice, William G.  
/ APPLICANT: Yang, Wengang  
/ TITLE OF INVENTION: Method for identifying compounds which inhibit binding of  
/ TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA  
/ FILE REFERENCE: 111021.143 (ACH-US1)  
/ CURRENT APPLICATION NUMBER: US/10/339,217  
/ CURRENT FILING DATE: 2003-01-09  
/ PRIOR APPLICATION NUMBER: US 60/347,369  
/ PRIOR FILING DATE: 2002-01-11  
/ NUMBER OF SEQ ID NOS: 144  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 108  
/ LENGTH: 496  
/ TYPE: PRF  
/ ORGANISM: HIV-1  
US-10-339-217-108

## Alignment Scores:

Pred. No.: 1,97e-117 Length: 496  
Score: 2382.00 Matches: 446  
Percent Similarity: 94.77% Conservative: 25  
Best Local Similarity: 89.74% Mismatches: 20  
Query Match: 85.93% Indels: 6  
DB: 14 Gaps: 2

US-09-475-704a-3 (1-1479) x US-10-339-217-108 (1-496)

QY 1 ATGGGCGCGCGCGCGCATCTCGGCGGCGGCGCAAGCTGGAGCGCGCATTCGCGC 60  
Db 1 MeGIyAlaArGAlaSeRlIeUaRGIyGIyLysLeuAsPlYsTrPlYInLysIleArG 20  
QY 61 CTGGCGCGCGCGCGCAAGAGTGCTACATGATGAGAGCACTGTGTGGCGCGCGAG 120  
Db 21 LeuArGIProGIyGIyLysLysLysTyRMeTLeuLysHlIeAlaITrPlaSeRArGIIn 40

QY 121 CTGGAGAAATTGCGCTTGAACCCCGGCTGCTGGAGACACCGAGGAGGCTGCAGAGATC 180  
Db 41 LeuGIuArGIyPheAlaAlaSeRlIeUaSeRGIyLeuGIuGIuTrSeRAsPGIyCySlYsGIInIle 60  
QY 181 ATCCGCGAGCTGCACCCCGCGCTGCGAGCCGCGAGCGAGAGCTGAGAGGCTGTTCAC 240  
Db 61 IIGInGIInLeuGIInProAlaLeuLysThrGIyInGIuGIuLeuArGISeRLeuTyRsn 80  
QY 241 ACCGTGACCACTTACTGCTGCGTGCAAGAGATCGAGTCCGCGACCAAGAGAGCGC 300  
Db 81 ThrValAlaTrnLeuTyCyEValHlIsAnSnIIGIuIIGIeArGIAsPThrIySGIuAla 100  
QY 301 CTGGAACAAATCGAGAGAGAGAGAGCAAGTGC-----CAGCAAGAAATC 345  
Db 101 LeuAsPArGIIGIuGIuGIuGIuLysCySGInGIInLysThrGIInGIInLysThr 120  
QY 346 CAGCGGCGCGAGCGCGCGCAAGAGGCGAAGTGAGCGCAAGCTACCCATGCTGAGAGAC 405  
Db 121 GIInGIInValGIuAlaAlaAsP--GIyLysValSeRGIInArGIyProIIGIeAlaGIInSn 139  
QY 406 CTGCAGAGCGCAGATGTGTCACACGACCATCAAGCCCGCCAGCTGAGAGCGCTGAGAG 465  
Db 140 LeuGIInGIuGIInMeTValHlIeGIInSeRLeuSeRProArGIInLeuAsnAlaITrValIlyS 159  
QY 466 GTGATCGAGAGAGAGGCTTCAGCCCGAGAGTATCCCATGTTCAAGCCCTGAGCGAG 525  
Db 160 ValIIGIuGIuLysAlaPheSeRProGIuIIGIeIIGIePrometPheThAlaLeuSeRGIIn 179  
QY 526 GGGCGCACCCCGCAGAGACTGACAGAGATGTTGAAACCGGTGGGCGGCACAGGCGCGC 585  
Db 180 GIyAlaArGIProGIInAsPLeuAsnThrMeTLeuAsnTrnValGIyGIyHlIeGIInAlaAla 199  
QY 586 ATGCAGATGTCTGAAGAGACCATCAACAGAGAGCGCGAGTGGAGACCGGCTGAGACCCC 645  
Db 200 MeTGIInMeTLeuLysArGIyThrIIGIeSnGIuGIuAlaAlaGIuITrPaRGIyLeuHlIePro 219  
QY 646 GTGCAGCGCGGCGCCATGCGCCCGCGCAAGTGCAGAGCCCGCGCGAGCAATCGCC 705  
Db 220 ValHlIsAlaGIyProValAlaProGIyIGInMeTArGIuProArGIySeRAsPlIeAla 239  
QY 706 GGCACACAGAGACCCCTGAGAGAGAGATGCCCTGGAGAGACAGACACCCCGCATTCGCC 765  
Db 240 GIyInTrnTrSeRAsnLeuGIInGIuGIInIIGIeSnTrpMeTThrAlaAsnProProIIGIePro 259  
QY 766 GTGGCGCAGATCTCAAGCGGTGATCATCTCGGCGCTGAACAAAGATGTGCGATGTAC 825  
Db 260 ValGIyAsPlIeTyRGIyArGIyTrPlIeIIGIeUGIyLeuAsnLysIIGIeAlaArGIeTyR 279  
QY 826 AGCCCGGTGAGCATCTGACATCAAGCAAGAGGAGGCGCCCAAGAGCCCTTCGCGACTAGTG 885  
Db 280 SerProValSeRlIeLeuAsPlIeLysGIInGIyProLysGIuProPheArGIyArGIyVal 299  
QY 886 GACCGCTTCCTCAAGACCTTCGCGCGAGAGAGACACCGAGAGTGAAGAATCGATG 945  
Db 300 AsPArGIyPhePheLysThrLeuArGIAlaGIuGIInAlaTrnGIInAsPValLysAsnITrPMeT 319  
QY 946 ACCGACACCTGCTGAGTGAAGAGCGCAACCCCGCATCTGCAAGACCATCTCGCGCTTC 1005  
Db 320 ThrAsPTrnLeuLeuValGIInAsnAlaAsnProAsPcySlYsThrIIGIeUArGIAlaLeu 339  
QY 1006 GGGCGCGCGCGCGCGAGAGAGTGAACCGCTTCGCAAGGCGGTGAGGCGCGCCAGC 1065  
Db 340 GIyProGIyAlaTrnLeuGIuGIuMeTMeThrAlaCySGInGIyValGIyLProSeR 359  
QY 1066 CACAAGCGCGCGGTGCTGCGCGAGCGAGTGAAGCGCAACACCAAGCGCTGATGAG 1125  
Db 360 HlIsLysAlaArGValLeuAlaGIuAlaMeTSeRGIInThraSeRnASeRlIeMeTMeTGIIn 379  
QY 1126 AAGAGCAACTTCAGAGGCGCGCGCGCATCTGCAAGTGTTCACATGCGCGAGAGAGGC 1185  
Db 380 AsnSeRAnPheLysGIySeRArGIyGIyIIGIeValLysCySPheAnCySGIlyLysValGIy 399  
QY 1186 CACATGCCCGCACTGCGCGCGCGCGCGCAAGAGGCGTCTGGAAGTGGCGGAGAG 1245

```

Db      400 HistidAlaArgAsnCyseAlaProArgLysGlyCysTrpLysCysGlyLysGlu 419
QY      1246 GGCACCAAGATGAGACTGCACCGCCGACGACCACTTCTGGGCAAGATCTGGCCC 1305
Db      420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyArgIleTrpPro 439
QY      1306 AGCCACAAGGCGCGCCCGCAACTTCTGCAGAGCGCGCCGACCGCCCGCCCGCC 1365
Db      440 SerHisLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
QY      1366 GCCGAGAGCTTCGCTTCGAGAGACCAACCCCGCCGACGAGCAGAGAGAGAGACCGC 1425
Db      460 AlaGlnSerPheArgPheGlnGlnIleThrProValProLysGlnGluProLysAspArg 479
QY      1426 GAGACCTGACCAAGCTTGAAAGAGCTGTTCGGCAACGACCCCTTGAGCCAG 1476
Db      480 GluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeuSerGln 496

```

Search completed: September 1, 2005, 23:50:51  
 Job time : 399.412 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 20:59:32; Search time 266.651 Seconds  
(without alignments)  
9259.839 Million cell updates/sec

Title: US-09-475-704A-4

Perfect score: 1509

Sequence: 1 atggggcggccggccagcatc.....ggcgccctgagccagcraa 1509

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1201	79.6	1515	4	US-09-475-515-4
2	1201	79.6	4472	4	US-09-475-515-75
3	1201	79.6	4608	4	US-09-475-515-76
4	1201	79.6	4689	4	US-09-475-515-74
5	1201	79.6	4766	4	US-09-475-515-73
6	1199	79.5	2031	4	US-09-475-515-7
7	1117.8	74.1	1853	4	US-09-475-515-5
8	1117.8	74.1	1865	4	US-09-475-515-78
9	1117.8	74.1	1865	4	US-09-475-515-79
10	1117.8	74.1	4319	4	US-09-475-515-6
11	1049.2	69.5	1268	4	US-09-475-515-9
12	971.4	64.4	1521	4	US-09-952-060-27
13	971.4	64.4	1532	4	US-09-974-702-1
14	971.4	64.4	1532	4	US-09-818-443-1
15	971.4	64.4	37474	4	US-09-952-060-25
16	968.4	64.2	4053	4	US-09-952-060-34
17	930.2	61.6	4307	4	US-09-936-572-2
18	929.4	61.6	4307	4	US-09-952-950-2
19	925.4	61.3	4327	4	US-09-936-572-14
20	925.4	61.3	4353	4	US-09-936-572-13
21	925.4	61.3	4642	4	US-09-936-572-12
22	925.4	61.3	9772	4	US-09-552-950-5
23	916.6	60.7	1482	4	US-09-818-443-4
24	913.6	60.5	1479	4	US-09-952-060-32
25	736.6	48.8	8366	4	US-09-872-733A-6
26	734.2	48.7	1496	4	US-09-184-418C-82
27	734.2	48.7	8972	4	US-09-184-418C-9

28	730.4	48.4	1476	4	US-09-991-258-4	Sequence 4, Appli
29	730.4	48.4	12523	4	US-09-991-258-1	Sequence 1, Appli
30	723.8	48.0	4338	4	US-09-872-733A-1	Sequence 1, Appli
31	712.8	47.2	9913	4	US-09-827-668-11	Sequence 11, Appli
32	709.6	47.0	1485	4	US-09-184-418C-73	Sequence 73, Appli
33	709.6	47.0	9010	4	US-09-184-418C-8	Sequence 8, Appli
34	708	46.9	1476	4	US-09-184-418C-100	Sequence 100, App
35	708	46.9	8959	4	US-09-184-418C-11	Sequence 11, Appli
36	663.6	44.0	1486	4	US-09-184-418C-38	Sequence 38, Appli
37	663.6	44.0	8992	4	US-09-184-418C-4	Sequence 4, Appli
38	651.8	43.2	8968	4	US-09-184-418C-1	Sequence 1, Appli
39	649.4	43.0	1479	4	US-09-184-418C-12	Sequence 12, Appli
40	642.2	42.6	7399	2	US-08-418-848A-9	Sequence 9, Appli
41	642.2	42.6	9709	2	US-08-188-583-5	Sequence 5, Appli
42	642.2	42.6	9709	3	US-08-388-353-1	Sequence 3, Appli
43	642.2	42.6	9709	3	US-08-488-551B-1	Sequence 1, Appli
44	642.2	42.6	9709	3	US-09-309-572-15	Sequence 15, Appli
45	642.2	42.6	9709	4	US-09-718-096-15	Sequence 15, Appli

## ALIGNMENTS

RESULT 1  
US-09-475-515-4  
Sequence 4, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: Patencin Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1515  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-475-515-4

Query Match 79.6%; Score 1201; DB 4; Length 1515;  
Best Local Similarity 88.5%; Pred. No. 2.8e-173;  
Matches 151; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY	1	ATGGGCGCGCGCGCAGCATCTCGCGCGCGAGAGCTGACCAAGTGGAGAAATCCGC	60
DB	7	ATGGGCGCGCGCGCAGCGTGTAGCGCGCGCAGCTGACCAAGTGGAGAAATCCGC	66
QY	61	CTGGCGCGCGCGCGAGAGCACTACTGTAAGCACTGTGTGGCCAGCCCGCAG	120
DB	67	CTGGCGCGCGCGCGAGAGCACTACTGTAAGCACTGTGTGGCCAGCCCGCAG	126
QY	121	CTGGAGGCGCTTGCCTGAACCCGCGCTGTGAGACCGCGAGGGCTGCAAGCAGATC	180
DB	127	CTGGAGGCGCTTGCCTGAACCCGCGCTGTGAGACCGCGAGGGCTGCGCCAGATC	186
QY	181	ATGAAGCACTGTCAGACCCCGCTGCGACCGCGACCGAGAGCTGCGCAGCTGTACAC	240
DB	187	CTGGCGCGAGCTGTGACCCAGCGCTGCGACCGCGAGCGAGAGCTGCGCAGCTGTACAC	246

241 ACCGTGGCCACCCTGTACTGCGTGACGCGCGATCGAGGTCCGACACCAAGAGAGCC 300  
 247 ACCGTGGCCACCCTGTACTGCGTGACGCGCGATCGAGGTCCGACACCAAGAGAGCC 306  
 301 CTGGACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 307 CTGGAGAAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366  
 361 GCGGACGAGCA-----AGGTAGGCAAGACTAACCCCACTCGTGAGAGAGAGAG 405  
 367 GCGGACGAGCAACCGGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426  
 406 CAGGGCAGATGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465  
 427 CAGGGCAGATGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486  
 466 ATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525  
 487 GTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546  
 526 GGCACCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585  
 547 GGCACCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606  
 586 CAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 645  
 607 CAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666  
 646 CAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705  
 667 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726  
 706 GGCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765  
 727 ACCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786  
 766 GGCAGATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 825  
 787 GGCAGATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846  
 826 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 885  
 847 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 906  
 886 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 945  
 907 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 966  
 946 GAGACCTGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 1005  
 967 GAGACCTGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 1026  
 1006 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 1065  
 1027 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 1086  
 1066 AAGGCGCGGTGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 1122  
 1087 AAGGCGCGGTGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 1146  
 1123 CAGAGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 1182  
 1147 CAGAGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 1206  
 1183 GGCACATGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 1242  
 1207 GGCACATGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 1266  
 1243 GAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1302  
 1267 GAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326  
 1303 CCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1362

1327 CCCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1368  
 1363 ACCGTGGCCACCCTGTACTGCGTGACGCGCGATCGAGGTCCGACACCAAGAGAGAG 1422  
 1369 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1428  
 1423 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1482  
 1429 ACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1488  
 1483 TTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1509  
 1489 TTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515

RESULT 2  
 US-09-475-515-75  
 ; Sequence 75, Application US/09475515A  
 ; Patent No. 6602705  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BARNETT, Susan  
 ; APPLICANT: ZUR MEGEDE, Jan  
 ; APPLICANT: SRIVASTAVA, Indresh  
 ; APPLICANT: LIAN, Ying  
 ; APPLICANT: HARTOG, Karin  
 ; APPLICANT: LIU, Hong  
 ; APPLICANT: GREER, Catherine  
 ; APPLICANT: SELBY, Mark  
 ; APPLICANT: WALKER, Christopher  
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
 ; FILE OF INVENTION: OF VIRUS-LIKE PARTICLES  
 ; FILE REFERENCE: 1621.002  
 ; CURRENT APPLICATION NUMBER: US/09/475.515A  
 ; CURRENT FILING DATE: 1999-12-30  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 75  
 ; LENGTH: 4472  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:  
 ; OTHER INFORMATION: gp160.modUS4.delV1/V2.gag.modSP2  
 US-09-475-515-75

Query Match 79.6%; Score 1201; DB 4; Length 4472;  
 Best Local Similarity 88.5%; Pred. No. 2.8e-173;  
 Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

1 AAGGCGCGGTGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 60  
 2939 AAGGCGCGGTGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 2998  
 61 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 120  
 2999 CCGGTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 3058  
 121 CTGGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 180  
 3059 CTGGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 3118  
 181 ATGAAGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 240  
 3119 CTGGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGATCTGAG 3178  
 241 ACCGTGGCCACCCTGTACTGCGTGACGCGCGATCGAGGTCCGACACCAAGAGAGAGAG 300  
 3179 ACCGTGGCCACCCTGTACTGCGTGACGCGCGATCGAGGTCCGACACCAAGAGAGAGAG 3238  
 301 CTGGACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 3239 CTGGACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3298

```
QY 361 GCCAGGGCA-----AGTGAAGCAGAACTACCCCACTCGTGAGAACTCTG 405
Db 3299 GCCGCGGCAACGGGCAACAGCAACCAAGTGAAGCAAGAACTACCTCGTGAGAACTCTG 3358
QY 406 CAGGGCCAGATGGTGAACAGGACCATCAGCCCCCGACCTGTAACCTCGTGAGAGTG 465
Db 3359 CAGGGCCAGATGGTGAACAGGACCATCAGCCCCCGACCTGTAACCTCGTGAGAGTG 3418
QY 466 ATCAGAGAAAGGCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCG 525
Db 3419 GTGAGAGAAAGGCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCG 3478
QY 526 GCCACCCCGCAGAGACTTGAAACAGATGTTGAACAACCTGCGGCGCCACACGAGCGCCATG 585
Db 3479 GCCACCCCGCAGAGACTTGAAACAGATGTTGAACAACCTGCGGCGCCACACGAGCGCCATG 3538
QY 586 CAGATGCTGAAGACACATCAACAGAGAGGCGCGGAGTGGGACCGGCTGACACCCCGTG 645
Db 3539 CAGATGCTGAAGAGACATCAACAGAGAGGCGCGGAGTGGGACCGGCTGACACCCCGTG 3598
QY 646 CAGGCGGCGCGCGTGGCCCCCGGACAGATGCGGACCCCCCGGAGAGCATGCGCGCG 705
Db 3599 CAGGCGGCGCGCGTGGCCCCCGGACAGATGCGGACCCCCCGGAGAGCATGCGCGCG 3658
QY 706 GCCACAGAGACCTTGAGAGAGAGATGCTGTGATGACCAAGCAACCCCGCTGCGCGTG 765
Db 3659 ACCACAGAGACCTTGAGAGAGAGATGCTGTGATGACCAACCCCGCTGCGCGTG 3718
QY 766 GGGCAGATCTACAGAGGGTGAATCATCTGAGGCGCTGAACAGATCGTGCGGATGTACAG 825
Db 3719 GGGCAGATCTACAGAGGGTGAATCATCTGAGGCGCTGAACAGATCGTGCGGATGTACAG 3778
QY 826 CCGGTGAGCATCTTGAGCATTCGCCAGGCGCCCAAGAGCCCTTCGCGACTACGTTGAG 885
Db 3779 CCGACAGAGCATCTTGAGCATTCGCCAGGCGCCCAAGAGCCCTTCGCGACTACGTTGAG 3838
QY 886 CGCTTTTCAAGACCTTGCGCGCGCGAGAGCCACCAAGAGCTGTGAAGAACTGTGATGAC 945
Db 3839 CGCTTTTCAAGACCTTGCGCGCGCTGAGAGAGCCAGAGCACTGTGAAGAACTGTGATGAC 3898
QY 946 GAGACCTGTGTGAGAGAGCCCAACCCCGACTGCAAGCACTCTGCGCGCTCTCGGC 1005
Db 3899 GAGACCTGTGTGAGAGAGCCCAACCCCGACTGCAAGCACTCTGCGCGCTCTCGGC 3958
QY 1006 CCGCGGCGCAACCTTGAGAGAGATGATGACCGCTGCGAGGCGCTGCGCGCGCCGCGCAC 1065
Db 3959 CCGCGGCGCAACCTTGAGAGAGATGATGACCGCTGCGAGGCGCTGCGCGCGCCGCGCAC 4018
QY 1066 AAGGCGCGGTGTGCGCGGCGAGTGAAGCCAG---CCAACAGCGTGAACATCATGATG 1122
Db 4019 AAGGCGCGGTGTGCGCGGCGAGTGAAGCCAGTGAACCCGCGCACTCATGATGATG 4078
QY 1123 CAGAAAGCAACTTCAAGGGCCCCCGGCGCAAGTCAAGTCACTCACTGCGCGCAAGAG 1182
Db 4079 CAGCGGGGCACTTCCGCAACCAAGCGAAAGCCGTCAAGTCACTCACTGCGCGCAAGAG 4138
QY 1183 GGGCAGATGCGCAAGAACTGCGCGCGCCCCCGCAAGAGGCTGTGGAAGTTCGCGCAAG 1242
Db 4139 GGGCAGACCGCGCAAGAACTGCGCGCGCCCCCGCAAGAGGCTGTGCGCGCGCGCG 4198
QY 1243 GAGGGCAGACATGTAAGAGCTGACACGAGCGCAGGCGCACTTCTGCGCGCAAGATG 1302
Db 4199 GAGGGCAGACATGTAAGAGCTGACACGAGCGCAGGCGCACTTCTGCGCGCAAGATG 4258
QY 1303 CCGAGCAGCAAGAGCGCGCGCGCACTTCTGCGCAAGCGAGCGAGCGCGCGCGCG 1362
Db 4259 CCGAGCTTCAAGAGCGCGCGCGCACTTCTGCGAGCGG-----C 4300
QY 1363 ACCGTGCGCAACCGCGCGCGCGAGAGCTTCGCTTGAAGAGACCAACCCCGCGCGCG 1422
Db 4301 CCGAGCGCGCAACCGCGCGCGCGAGAGAGCTTCGCTTGAAGAGAGAAACCAACCCCG 4360
```

```
QY 1423 AAGCAGAGGCCCAAGAGACCGGAGGCCCTTACCGGAGCCCTTGACCGCCCTGCGAGCCTG 1482
Db 4361 AGCCAGAGACAGAGACCCATTCAGACAGAGCTGTACCTCCCTGACCAAGCCTGCGAGCCTG 4420
QY 1483 TTCGCGAGCGCGCCCTTGAGCCAGTAA 1509
Db 4421 TTCGCGAGAGACCCAGAGCCAGTAA 4447

RESULT 3
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 4608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-76

Query Match          79.6%; Score 1201; DB 4; Length 4608;
Best Local Similarity 88.5%; Pred. No. 2,8e-173;
Matches 1551; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY 1 ATGGCGCGCGCGCGCAGCATCTGCGCGCGAGAAAGCTGGAACAAGTGGAGAAATCCGC 60
Db 3075 ATGGCGCGCGCGCGCAGCGTGTGAGCGGCGGAGCTGGAACAAGTGGAGAAATCCGC 3134
QY 61 CCGCGCGCGCGCGCGCAAGAAAGCACTACATGCTGAAGCACTGTGTGGCGCAGCGCGAG 120
Db 3135 CCGCGCGCGCGCGCGCAAGAAAGTAAAGCTGAAGCACTGTGTGGCGCAGCGCGAG 3194
QY 121 CTGAGAGGCTTGCGCGCTGAACCCCGGCTGTGAGAACCGCGAGGAGGCTGCAAGCAGATC 180
Db 3195 CTGAGAGGCTTGCGCGCTGAACCCCGGCTGTGAGAACCAAGGAGGCTGCGCGCAAGTC 3254
QY 181 ATGAAGCAAGCTGACAGCCCGCTGCAACCGGCACTGAGAGAGCTGCGACGCTGTACAC 240
Db 3255 CTGAGCGAGCTGCAAGCCAGCGCTGCAAGCCGCAAGCGAGAGCTGCGACGCTGTACAC 3314
QY 241 ACCGTGGCGCAACCTGTACTGCGGTGCAAGCGCGCAATCGAGTCCGCGAGACCAAGAGAGCG 300
Db 3315 ACCGTGGCGCAACCTGTACTGCGGTGCAAGCGCGCAATCGAGTCCGCGAGAGAGCG 3374
QY 301 CTGGAACAAGATGAGAGAGAGAGAGAAAGTCCAGCAAGAGAGAGAGAGAGAGAGAGAG 360
Db 3375 CTGGAAGAAATGAGAGAGAGAGAGAAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 3434
QY 361 GCCAGCGCA-----AGTGAAGCAGAACTACCCCATGTTGTCGAGAACTCTG 405
Db 3435 GCCGCGCGCACCGGCAACAGAGCCAGGTGAGCCGAACTACCCCATGTTGTCGAGAACTCTG 3494
QY 406 CAGGAGCAGATGGTGAACAGGACCATCAGCCCCCGACCTGTAACGCTGTGAGTGAAGTG 465
```

Db 3495 CAGGGCCAGATGTGTGACCAAGGCATCAGCCCCGCACTGTAAGCTGTGGTGAAGTGTG 3554  
 QY 466 ATCGAGGAGAGGCTTTCAGCCCCGAGGTGATCCCATATTCAACGCGCTTGAGCGAGGGC 525  
 Db 3555 GTGGAGGAGAGGCTTTCAGCCCCGAGGTGATCCCATATTCAACGCGCTTGAGCGAGGGC 3614  
 QY 526 GCCACCCCCAGAGACTTGAACAAGATTTGAACACCGGTGGGGGAGCCAGAGCGCGCATG 585  
 Db 3615 GCCACCCCCAGAGACTTGAACAAGATTTGAACACCGGTGGGGGAGCCAGAGCGCGCATG 3674  
 QY 586 CAGATGCTGAAGACCACTCAACGAGAGGCGCGGAGTGGGACCGCTGCACCCCGTG 645  
 Db 3675 CAGATGCTGAAGAGCACTCAACGAGAGGCGCGGAGTGGGACCGCTGCACCCCGTG 3734  
 QY 646 CAGGCGGCGCGCGTGGCGCGCGGAGTGGCGGACCGCGCGGAGCGGACCTGCGCGGCG 705  
 Db 3735 CAGGCGGCGCGCGTGGCGCGCGGAGTGGCGGACCGCGCGGAGCGGACCTGCGCGGCG 3794  
 QY 706 GCCACCAAGACCTTGCAGAGAGAGATCGCTGTGATGACGACACCCCGCTGCGCGGCG 765  
 Db 3795 ACCACCAAGACCTTGCAGAGAGAGATCGCTGTGATGACGACACCCCGCTGCGCGGCG 3854  
 QY 766 GCGGACATCTCAACGCGGTGATCATCTGCGCTGTAACAGATCTGCGATGTACAGC 825  
 Db 3855 GCGGAGATCTCAACGCGGTGATCATCTGCGCTGTAACAGATCTGCGATGTACAGC 3914  
 QY 826 CCGGTGAGATCTTGGACATCTGCGCGGAGCGGCGCGGACCGCTGCGGACCTACGTTGAC 885  
 Db 3915 CCGACCAAGATCTTGGACATCTGCGCGGAGCGGCGCGGACCGCTGCGGACCTACGTTGAC 3974  
 QY 886 CCGCTTTCAGAGCCTTGCAGAGAGAGATGATGACCGGCTGCGAGGCGGCGCGCGCGAC 945  
 Db 3975 CCGTTTCAAGAGCCTTGCAGAGAGAGATGATGACCGGCTGCGAGGCGGCGCGCGCGAC 4034  
 QY 946 GAGACCTTGTGTGCAAAACCGCAACCCGACTGCAACCACTCTGCGCGCTCTCGGC 1005  
 Db 4035 GAGACCTTGTGTGCAAAACCGCAACCCGACTGCAACCACTCTGCGCGCTCTCGGC 4094  
 QY 1006 CCGGCGGCGACCTTGGAGAGATGATGACCGGCTGCGAGGCGGCGGCGCGCGCGAC 1065  
 Db 4095 CCGGCGGCGACCTTGGAGAGATGATGACCGGCTGCGAGGCGGCGGCGCGCGCGAC 4154  
 QY 1066 AAGGCGCGGTGTGCGAGGCGAGTGAAGCCAGG---CCAAAGCGCTGAACTCATGATG 1122  
 Db 4155 AAGGCGCGGTGTGCGAGGCGAGTGAAGCCAGTGAAGCCCGCGACATCATGATG 4214  
 QY 1122 CAGAAAGCACTTCAAGGCGCGCGCGGCAAGTGAAGTCTTCACTGCGCGCAAGAG 1182  
 Db 4215 CAGCGGCGCACTTCCGCAACGAGCGGAAGCCGTCMAAGTGTCTCAACTGCGCGCAAGAG 4274  
 QY 1183 GGCACATCGCCCAAGAACTGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGCAGAG 1242  
 Db 4275 GGCACATCGCCCAAGAACTGCGCGCGCGCGCGCGCAAGAGGCGTGTGTGCGCGCGC 4334  
 QY 1243 GAGGCGCACCAATGAAGACTGCAACGAGCGCAGGCCAATTCTCTGGGCAAGATTG 1302  
 Db 4335 GAGGCGCACCAATGAAGACTGCAACGAGCGCAGGCCAATTCTCTGGGCAAGATTG 4394  
 QY 1303 CCCAGGCAAGAGGCGCGCGCGCAATTCTGTGCAAGCCGACGAGCGCGCGCGCGC 1362  
 Db 4395 CCCAGGCAAGAGGCGCGCGCGCAATTCTGTGCAAGCGC-----C 4436  
 QY 1363 ACCGTGCGACGCGCGCGCGCGCGAGCTTCCGTTGAGAGAGCAACCCCGCGCGC 1422  
 Db 4437 CCGAGGCCACCGCGCGCGCGCGAGAGCTTCCGTTGAGAGAGCAACCCCGCGCGC 4496  
 QY 1423 AAGGAGAGCCCAAGAGACCGAGCGCTTACCGGAGCGCGCTGACCGCGCTGCGAGCGCTG 1482  
 Db 4497 AAGGAGAGCCCAAGAGACCGAGCTTGAACAAGAGCTGTACCCCTGTACAGCGCTGCGAGCGCTG 4556  
 QY 1483 TTTCGCGAGCGCGCGCTTGAAGCACTAA 1509  
 Db 4557 TTTCGCGAGCGCGCGCTTGAAGCACTAA 4583

RESULT 4  
 US-09-475-515-74  
 / Sequence 74, Application US/09475515A  
 / Patent No. 6602705  
 / GENERAL INFORMATION:  
 / APPLICANT: BARNETT, Susan  
 / APPLICANT: ZUR MEGEDE, Jan  
 / APPLICANT: SRIVASTAVA, Indresh  
 / APPLICANT: LIAN, Ying  
 / APPLICANT: LIU, Hong  
 / APPLICANT: GREER, Catherine  
 / APPLICANT: SELBY, Mark  
 / APPLICANT: WALKER, Christopher  
 / TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
 / FILE REFERENCE: 1621.002  
 / CURRENT APPLICATION NUMBER: US/09/475.515A  
 / NUMBER OF SEQ ID NOS: 90  
 / SOFTWARE: PatentIn Ver. 2.0  
 / SEQ ID NO 74  
 / LENGTH: 4689  
 / TYPE: DNA  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: Description of Artificial Sequence:  
 / OTHER INFORMATION: SPI60.modSF162.gag.modSF2  
 US-09-475-515-74  
 Query Match 79.6%; Score 1201; DB 4; Length 4689;  
 Best Local Similarity 88.5%; Pred. No. 2.8e-173; Indels 36; Gaps 3;  
 Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;  
 QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGAGAGAGTGTGACCAAGTGGAGAAATCCGC 60  
 Db 3156 ATGGGCGCGCGCGCGAGCGTGTGAGCGCGCGCGAGCGTGTGAGCAAGTGGAGAAATCCGC 3215  
 QY 61 CTGCGCGCGCGCGCGAGAGCACTACTGTGAAGCACTGTGTGTGGCCAGCCCGAG 120  
 Db 3216 CTGCGCGCGCGCGCGAGAGCACTACTGTGAAGCACTGTGTGTGGCCAGCCCGAG 3275  
 QY 121 CTGAGGCGCTTGCCTGTAACCCCGCGCTGTGAGAGCGCGGAGGCGTGTGAGAGATC 180  
 Db 3276 CTGAGGCGCTTGCCTGTAACCCCGCGCTGTGAGAGCGCGGAGGCGTGTGAGATC 3335  
 QY 181 ATGAAGCACTGCAAGCCCGCGCTGTGAGAGCGGCAACGAGAGTGTGCGAGCTGTACAC 240  
 Db 3336 CTGAGGCGCTGCAAGCCCGCGCTGTGAGAGCGGCAACGAGAGTGTGCGAGCTGTACAC 3395  
 QY 241 ACCGTGCGACCTGTATCTGTGTGACGCGCGCATCGAGTGTCCGAGACCAAGAGGCC 300  
 Db 3396 ACCGTGCGACCTGTATCTGTGTGACGCGCGCATCGAGTGTCCGAGACCAAGAGGCC 3455  
 QY 301 CTGGAAGATCTGAGAGAGAGAGCAAGATCCAGAGAGAGAGAGAGAGAGAGAG 360  
 Db 3456 CTGGAAGATCTGAGAGAGAGAGCAAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 3515  
 QY 361 GCCGACGCA-----AGGTGACCAAGATCACTCCCATGTGCAGAACTTG 405  
 Db 3516 GCCGACGCA-----AGGTGACCAAGATCACTCCCATGTGCAGAACTTG 3575  
 QY 406 CAGGCGCAGATGTGTGACCAAGGCAATCAGCCCGCGCACTGTAAAGCTGTGAGTGTG 465  
 Db 3576 CAGGCGCAGATGTGTGACCAAGGCAATCAGCCCGCGCACTGTAAAGCTGTGAGTGTG 3635  
 QY 466 ATCGAGGAGAGGCGCTTACGCGCGAGGTGATCCCAATGTTTCAACCGCGCTGAGAGAGGCG 525  
 Db 3636 GTGGAGGAGAGGCGCTTACGCGCGAGGTGATCCCAATGTTTCAACCGCGCTTACGAGAGGCG 3695  
 QY 526 GCCACCCCGCAGGACTGTGAACAGATGTTGAACACCGTGGCGGCGACCAAGCGCGCATG 585

```

Db      3696 GCCACCCCCAGAGACCGTGAACAGATGTTGAACACCTGTGGCCGCCACAGAGCCGCCCATG 3755
Qy      586 CAGATGCTGAAGACACCATCAACGAGAGAGCCGCCAGTGGAGCCGCTGCACCCCGTG 645
Db      3756 CAGATGCTGAAGAGACCATCAACGAGAGAGCCGCCAGTGGAGCCGCCGTGCACCCCGTG 3815
Qy      646 CAGGCGGCGCCCGTGGCCCGCCGAGTGGAGCCGCCCGGCGGCGGACATCGCCGCG 705
Db      3816 CAGCGCGGCGCCCATCGCCCGCCGAGATGCGGAGCCCGCGGCGGCGGACATCGCCGCG 3875
Qy      706 GCCCAGGAGACCTGTGAGAGAGCAGATCGCTGTGATGACGACCAACCCCGCGCCGCG 765
Db      3876 ACCACAGAGACCTGTGAGAGAGAGATCGGTGTGATGACCAACAACCCCGCATCCCGCG 3935
Qy      766 GCGGACATCTACAGCGGTGTGATCATCTGGGCTTGAACAGATCGTGGATGTACAC 825
Db      3936 GCGGAGATCTACAGCGGTGTGATCATCTGGGCTTGAACAGATCGTGGATGTACAC 3995
Qy      826 CCCGTGAGCATCTGTGACATCGGCGGAGCCCAAGAGCCCTTCCGCGATCACTGTGAC 885
Db      3996 CCCACAGACATCTGTGACATCGGCGGAGCCCAAGAGCCCTTCCGCGATCACTGTGAC 4055
Qy      886 CGCTTCTTAAGACCTGTGCGCGCCGAGCAGGCCACCAAGACGTGAAGAACTGTGATGAC 945
Db      4056 CGCTTCTTAAGACCTGTGCGCGCCGAGCAGGCCACCAAGACGTGAAGAACTGTGATGAC 4115
Qy      946 GAGACCTGTGTGTGAGAAACGCGCAACCCGACTGTGAAGACATCTCGCGCTCTCGAC 1005
Db      4116 GAGACCTGTGTGTGAGAAACGCGCAACCCGACTGTGAAGACATCTCGCGCTCTCGAC 4175
Qy      1006 CCCGCGGCGACCTGTGAGAGATGATGACCGCTGTGACAGGCGGTGGCGCCCGCGCAC 1065
Db      4176 CCCGCGGCGACCTGTGAGAGATGATGACCGCTGTGACAGGCGGTGGCGCCCGCGCAC 4235
Qy      1066 AAGGCGCGGCGTGTGCGCGAGGCGATGAGCAGG---CCAAACGCGTGAACATCATGATG 1122
Db      4236 AAGGCGCGGCGTGTGCGCGAGGCGATGAGCAGGCGCGCGACATCATGATG 4295
Qy      1123 CAGAGAGCACTTCAAGGCGCCCGCGCGCAACGTCAGTGTCTTCAACTGCGGCAAGAG 1182
Db      4296 CAGCGCGGCACTTCCGCAACAGCGGAGACCGTCAAGTGTCTTCAACTGCGGCAAGAG 4355
Qy      1183 GCGCAATGCGCAAGAACTGCGCGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAAGTCCGCG 1242
Db      4356 GCGCAATGCGCAAGAACTGCGCGCGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAAGTCCGCG 4415
Qy      1243 GAGGCGCAACAGATGAAAGACTGCAACGAGCGCGCAAGCCATTCTGTGGCAAGATCTGG 1302
Db      4416 GAGGCGCAACAGATGAAAGACTGCAACGAGCGCGCAAGCCATTCTGTGGCAAGATCTGG 4475
Qy      1303 CCCAGCGCAAGAGGCGCGCGCGCAACTTCTGTGAGAACCGGAGCGAGCCCGCGCGCC 1362
Db      4476 CCCAGCGCAAGAGGCGCGCGCGCAACTTCTGTGAGAGCG-----C 4517
Qy      1363 ACCGTGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1422
Db      4518 CCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4577
Qy      1423 AAGCGAGAGCGCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1482
Db      4578 AAGCGAGAGCGCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4637
Qy      1483 TTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509
Db      4638 TTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4664

```

RESULT 5  
 US-09-475-515-73  
 ; Sequence 73, Application US/09475515A  
 ; Patent No. 6602705  
 ; GENERAL INFORMATION:

```

; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
US-09-475-515-73

```

```

Query Match      79.6%; Score 1201; DB 4; Length 4766;
Best Local Similarity 88.5%; Pred. No. 2.8e-173;
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

```

```

Qy      1 ATGGAGCGCGCGCGCGAGATCTGTGCGCGGAGAGCGTGAAGCAAGTGGAGAAAGATCCGC 60
Db      3223 ATGGAGCGCGCGCGCGAGGCTGTGAGCGCGCGCGGCTGTGAACAAATGGAGAAAGATCCGC 3292
Qy      61 CTGCGCGCGCGCGCAAGACATCACTGTGAAGCACTGTGTGGCGCAGCCCGCAG 120
Db      3223 CTGCGCGCGCGCGCAAGAAAGTACAGCTGAAGACATCTGTGTGGCGCAGCCCGCAG 3352
Qy      121 CTGAGGCGCTTGCCCTTGAACCCCGCGCTGTGTGAGACCGCGGAGGCTGTGAAGCATC 180
Db      3353 CTGAGGCGCTTGCCCTTGAACCCCGCGCTGTGTGAGACCAAGGAGGCTGTGCGCAGATC 3412
Qy      181 ATGAAGCAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db      3413 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3472
Qy      241 ACCGTGCGCAACCTGTGATCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      3473 ACCGTGCGCAACCTGTGATCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3532
Qy      301 CTGGAAGAGTGAAGAGAGAGAGAGCAAGTCCAGCAAGAAAGCCAGAGCGCGCAAGAG 360
Db      3533 CTGGAAGAGTGAAGAGAGAGAGAGCAAGTCCAGCAAGAAAGCCAGAGCGCGCGCGCG 3592
Qy      361 GCCGAGCGCGCA-----AGGTGAGCGCAAGAACTTACCCCATGTGTGCAAGAACTTG 405
Db      3593 GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3652
Qy      406 CAGGCGCGAGTGTGCAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
Db      3653 CAGGCGCGAGTGTGCAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3712
Qy      466 ATCGAGAGAAAGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 525
Db      3713 GTGAGAGAGAAAGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3772
Qy      526 GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 585
Db      3773 GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3832
Qy      586 CAGATGCTGAAGAGACACCATCAACGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 645
Db      3833 CAGATGCTGAAGAGAGACCATCAACGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3892

```

QY	646	CAGGCGGGCCCTGGGCCCCCGGCGAGATGGCGACCCCGGCGAGGACATGCGCGG	705
Db	3893	CAGCGGGGCCCATATGCCCCCGGCGAGATGGCGAGCCCCCGGCGAGCATGCGCGG	3952
QY	706	GCCACCAACACCTTGCAGAGGACGATTCGCTGGATGACACAGCAACCCCGTGCCGTG	765
Db	3953	ACCAACAGACACCTTGCAGAGGACGATTCGCTGGATGACCAACCAACCCCGCATCCCGTG	4012
QY	766	GGCCACATCTACAAAGCGGTGATCATCTTGGGGCTTGAAACAGATGTGCGGATGTACAGC	825
Db	4013	GGCGAGATCTACAAAGCGGTGATCATCTTGGGGCTTGAAACAGATGTGCGGATGTACAGC	4072
QY	826	CCCGTAGACATCTTGGACATCCGCGAGAGGCCCAAGAGACCTTTCGCGACTACGTGAC	885
Db	4073	CCCAACACATCTTGGACATCCGCGAGAGGCCCAAGAGACCTTTCGCGACTACGTGAC	4133
QY	886	CGCTTCTTCAAGACCTTGGCGCGCGAGCAGCGCACCCAGAGCTGAAGATCTGATGACC	945
Db	4133	CGCTTCTTCAAGACCTTGGCGCGCTTGAGAGCGCACCGTAAACATCTGATGTACC	4192
QY	946	GAGACCTCGCTGGTGCAGAAAGCCCAACCCGACTGCAAGACCATCTGCGGCTTGCAG	1009
Db	4193	GAGACCTCGCTGGTGCAGAAAGCCCAACCCGACTGCAAGACCATCTTGAAGCTTTCGAG	4255
QY	1006	CCCGGCGCCACCCTTGAGAGATGATGACCGGCTTGCAGAGGCGTGGGCGGCCCGGCAC	1065
Db	4253	CCCGGCGCCACCCTTGAGAGATGATGACCGGCTTGCAGAGGCGTGGGCGGCCCGGCAC	4312
QY	1066	AAGGCCCCGCTGCTGGCGGAGGCGATGAGCCAG---CCAAACAGGTAAATCATATGATG	1122
Db	4313	AAGGCCCCGCTGCTGGCGGAGGCGATGAGCCAGGTAAACACCGGACATCATATGATG	4372
QY	1123	CAGAAGAGCACTTGAAGGCGCCCCCGGCGCAACGTCAAGTCTTCAACTGCGCAAGAG	1182
Db	4373	CAGCGCGGCACTTCCGCAACCAAGGGAAAGACGTCAAGTCTTCAACTGCGCAAGAG	4432
QY	1183	GGCCACATGCGCAAGACTGCGCGCGCCCCCGCAAGAAAGGCTGCTGGAATGCGCGAAG	1242
Db	4433	GGCCACACCGCCAGAGAACTGCGCGCGCCCCCGCAAGAAAGGCTGCTGCGCTGCGCGCG	4492
QY	1243	GAGGCGCAACCATATGAAGACTGACCCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGG	1302
Db	4493	GAGGCGCAACCATATGAAGACTGACCCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGG	4552
QY	1303	CCCAACCAACAAAGGCGCCCGCGCAACTTCTTGCAAAACCGACGAGCGCGCGCGCCC	1362
Db	4553	CCCACTCTCAAAAGGCGCGCCCGCGCAACTTCTTGCAAGCGG-----C	4592
QY	1363	ACCGTGGCCACCGCCCCCGCGCGAGAGCTTCGCTTCAAGAGACCAACCCCGCGCCCC	1422
Db	4595	CCCGAGCCCAACGCCCCCGCGAGAGACTTTCGCTTGGCGAGAGAGACCAACCCCC	4652
QY	1423	AAGCAGAGGCCCAAGAGACCGCAGCGCTTACCGCGAGCCCTTGACCGCTTGCAGCGCTG	1482
Db	4655	AACCAAGAGCAGAGAGCCATTCGACAAAGAGCGTGTACCCCTGACCAAGCCTTGCGAGCGTG	4712
QY	1483	TTTCGGCAGGGGCCCCCTGAGCCAGTAA	1509
Db	4715	TTTCGGCAAGACCCCAAGCCAGCAAGTAA	4741

```

RESULT 6
US-09-475-515-7
; Sequence 7, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEBE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine

```

```

1  APPLICANT: SELBY, Mark
2  APPLICANT: WALKER, Christopher
3  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
4  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
5  FILE REFERENCE: 1621.002
6  CURRENT APPLICATION NUMBER: US/09/475,515A
7  CURRENT FILING DATE: 1999-12-30
8  NUMBER OF SEQ ID NOS: 90
9  SOFTWARE: PatentIn Ver. 2.0
10 SEQ ID NO 7
11
12 LENGTH: 2031
13
14 TYPE: DNA
15
16 ORGANISM: Artificial Sequence
17
18 FEATURE:
19
20 OTHER INFORMATION: Description of Artificial Sequence: synthetic
21
22 OTHER INFORMATION: HIV-Gag/HCV-core fusion polypeptide
23
24 US-09-475-515-7

```

Query Match	79.5%	Score 1199;	DB 4;	Length 2031;
Best Local Similarity	88.5%;	Pred. No. 5.6e-173;		
Matches 1349;	Conservative	0;	Mismatches 140;	Indels 36; Gaps 3;

QY	1	TTGGGCGCGCCGCGCAGCATCTCTGCGCGGCAGAAAGCTGGACAAGTGGGAAGAAAGTCCG	60
Db	7	ATGGGCGCGCCGCGCAGCGTCTAGCGCGCGCGACGTGGACAACTGGGGAAGAAAGTCCGC	66
QY	61	CTGGCGCCCGCGCGCAAGAACCTACTATCTGTGAAGCACTGTGTGTGGGCAAGCCGCGAG	120
Db	67	CTGGCGCCCGCGCGCAAGAAAGTAACTAAAGCTGAAGCACTGTGTGTGGGCAAGCCGCGAG	126
QY	121	CTGGAGGGCTTTCGCTTGGAACTCCCGGCTGTGTGGAAACCGCGGAGGGCTTGAAAGAGATC	180
Db	127	CTGGAGGGCTTTCGCTTGGAACTCCCGGCTGTGTGGAGAACCGGAGGGCTTGCGCCAGATC	186
QY	181	ATGAAGCAGCTGCAGCCGCGCCTTGCAAGCCGCGACCCGAGAGCTTCGCGAGCTGTAAAC	240
Db	187	CTGGGCGACGTGCAGCCCGACCTTGCAAGCCGCGACGAGAGAGCTTCGCGAGCTGTAAAC	246
QY	241	ACCGTGGCCACCTGTACTGTGCTGCACGCGCGGACTGAGGTTCCGCGACCAAGAGAGCC	300
Db	247	ACCGTGGCCACCTGTACTGTGCTGCACGAGCGCATCGACTCAAGGACCAAGAGAGCC	306
QY	301	CTGGACAAAGATCGAGAGAGAGACAGAACAAAGTCCCGACGAGAAAGCCACACAGGCCAAAGAG	360
Db	307	CTGGAGAAAGATCGAGAGAGAGACAGAACAAAGTCCAAAGAAAGGCCCAACAGGCCCGCC	366
QY	361	GCCGACGGCA-----AGGTGAGCCGAACTACCCCATGTGTCAGAAACTG	405
Db	367	GCCGCGGCAACCGGCAACGACGACCGAGTGAAGCCAGAACTACCCCATGTGTCAAACTG	426
QY	406	CAGGCGCAGATGTGCACACAGGCGCATCAGCCCGCGCACCTTGAAAGCCTTGAGTGAAGGTG	465
Db	427	CAGGCGCAGATGTGTGCACACAGGCGCATCAGCCCGCGCACCTTGAAAGCCTTGAGTGAAGGTG	486
QY	466	ATCGAGGAAAGGCTTCAAGCCCGAGGTGATCCCATATTTCAACGCGCCTTGACGCAAGGC	525
Db	487	GTGGAGGAAAGGCTTCAAGCCCGAGGTGATCCCATATTTCAAGGCTTGACGCAAGGC	546
QY	526	GCCACCCCCCGAGACTGAAACAGATGTTTGAACACGTGGGCGGCGCACAGGCGCCCATG	585
Db	547	GCCACCCCCCGAGACTTGAACAGATGTTTGAACACCGTGGGCGGCGCACAGGCGCCCATG	606
QY	586	CAGATGTGAAGGACCAACCATCAACGAGAGAGCCGCGGAGTGGGAACGCGCTTGACCCCGTG	645
Db	607	CAGATGTGAAGGAGACCATCAACGAGAGAGCCGCGGAGTGGGAACGCGCTTGACCCCGTG	666
QY	646	CAGGCGGCGCGGTGGCCCCCGGCGCAATGGCGCAACCCCGCGGCAAGGACATTCGCGGC	705
Db	667	CAGGCGGCGCGCATTCGCCCGCGCGCAGATGGCGCGAGCCCGCGGCGAGGACATTCGCGGC	726
QY	706	GCCACACACACCTTCGAGAGAGAGATGGCTGTGAATGACAGCAACCCCGGTGCGGTG	765
Db	727	ACCAACACACCTTCGAGAGAGAGATGGCTGTGAATGACCAACCCCGCATTCGCGGTG	786





```

Db      847 CCCACGAGCATCTTGACATCCGCCAGGGCCCCCAGAGGCCCTTCCGCACTACGTGAC 906
Qy      886 CGCTTTCAAGACCTCTGGCGCCGAGCAGGCCACCCAGAGCGTGAAGAACTGGATGACC 945
Db      907 CGCTTTCAAGACCTCTGGCGCCGAGCAGGCCACCCAGAGCGTGAAGAACTGGATGACC 966
Qy      946 GAGACCTGCTGTGTGAGAACCGCCACCCCGACTGTGACAACTCTGTGCGCTCTCGAC 1005
Db      967 GAGACCTGCTGTGTGAGAACCGCCACCCCGACTGTGACAACTCTGTGAGGCTCTCGAC 1026
Qy      1006 CCCGCGCCACCTTGAGAGAGATGATGACCCCTGCGCAGAGGGCTGGCGCCCGCCGAC 1065
Db      1027 CCCGCGCCACCTTGAGAGAGATGATGACCCCTGCGCAGAGGGCTGGCGCCCGCCGAC 1086
Qy      1066 AAGGCCCGGTGTGTGCGCCGAGCGAGTATGACGAGG---CCACAGCGTGAACATCATGAG 1122
Db      1087 AAGGCCCGGTGTGTGCGCCGAGCGAGTATGACGAGG---CCACAGCGTGAACATCATGAG 1146
Qy      1123 CAGAAAGCACTTCAAGGGCCCCCGCCGCAACGTCAAGTGTCTCAACTGTGCGCAAGAG 1182
Db      1147 CAGCGCGCAACTTCCGCAACCGAGGAAAGACGTCAAGTGTCTCAACTGTGCGCAAGAG 1206
Qy      1183 GGCACATGCTCCCAAGACTGCGCGCCCGCCGCAAGAGGGCTGCTGGAATGCGGCAAG 1242
Db      1207 GGCACACCGCAGAACTGCGCGCCCGCCGCAAGAGGGCTGCTGCGCTGCGCGCCG 1266
Qy      1243 GAGGGCACAGATGAAGACTGACCGAGCGCCAGGCCCACTTCTGTGGGCAAGATCTGG 1302
Db      1267 GAAAGCACCAATGAAGATTGCACTGAGAGACAGGCTTAATTTTAAAGGAAGATCTGG 1326
Qy      1303 CCCAGCACAAAGGGCGCCCGCGCACTTCTGTGACAAACCGCAGCGAGCCCGCGCCG 1362
Db      1327 CTTTCTCAAGAGGAGAGCGCAGGGAATTTTCTTCAAGAGC-----A 1368
Qy      1363 ACCGTGCCCAACGCGCCCGCCGAGAGCTTCCGTTGAGAGAACACCCCGCGCCG 1422
Db      1369 CCAGACCCCAACGCGCCCGCCGAGAGCTTCAAGTTTGGGAGGAGAAACAACTCCG 1428
Qy      1423 AAGCAGAGACCCCAAGAGACCGGAGCCCTTACCGGAGACCCCTGTACCGCTGTGCGAGCTG 1482
Db      1429 TCTCAGAAACAGAGAGCGCGATAGCAAGAACTGTATCTTTAACTTCCCTCAGATCATCTC 1488
Qy      1483 TTCGCGAGCGGCCCTCTGAGCCAGTAA 1509
Db      1489 TTTGGCAACGACCCCTGTCAAGTAA 1515

```

RESULT 8  
US-09-475-515-78  
Sequence 78, Application US/09475515A  
Patent No. 6602705

GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR WEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475, 515A  
CURRENT FILING DATE: 1999-12-30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 78  
LENGTH: 1865  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: GPI  
US-09-475-515-78

```

Query Match      74.1%; Score 1117.8; DB 4; Length 1865;
Best Local Similarity 85.1%; Pred. No. 1e-160;
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

Qy      1 ATGGCGCGCCGCGCAGCATCTGCGCGGAGAGAGCTGGAACAAGTGGAGAAATCCGC 60
Db      13 ATGGCGCGCCGCGCAGCGTGTGAGCCGCGCGAGCTGGAACAAGTGGAGAAATCCGC 72
Qy      61 CTGCGCCCGCGCGCAAGAGCATCTGTAAGCACTGTGTGGGCCAGCCGCGAG 120
Db      73 CTGCGCCCGCGCGCAAGAAAGTAACAAGCTGAAGACATGTGTGGGCCAGCCGCGAG 132
Qy      121 CTGAGAGGCTTCCGCTTGAACCCCGGCTGTGTGAGAACCGCGAGGGCTGCAAGCATC 180
Db      133 CTGAGAGGCTTCCGCTTGAACCCCGGCTGTGTGAGAACCGAGGGCTGTGCGCGCATC 192
Qy      181 ATGAAGCAGCTGCAACCCGCTGTGAGACCGGACCGAGAGCTGCGCAGCTGTACAAC 240
Db      193 CTGGGCGCAGCTGCAACCCGCTGTGAGACCGGACCGAGAGCTGCGCAGCTGTACAAC 252
Qy      241 ACCGTGGCCACCTGTACTGTGTGACGCGCGCATCGAGTCCGCGACCAAGAGGCGC 300
Db      253 ACCGTGGCCACCTGTACTGTGTGACGCGCGCATCGAGTCCGCGACCAAGAGGCGC 312
Qy      301 CTGGAACAATCTGAGAGAGAGACAAACATGCCAGCAAGAAAGACCCAGAGCCAGAGAG 360
Db      313 CTGGAAGAATCTGAGAGAGAGACAAACATGCCAGCAAGAAAGGCGCCAGAGGCGCGC 372
Qy      361 GCCGAGCGCA-----AGGTGAGCCAGAACTACCCCATGTGTGAGAACTGTG 405
Db      373 GCCGCGCGCAACCGGCAACAGACAGCGAGTGAAGCCAGAACTACCCCATGTGTGAGAACTGTG 432
Qy      406 CAGGSCCAGATGTGTGACCAAGGCAATCAGCCCGGCACTGTGAACGCTGTGGGTGAAGTGTG 465
Db      433 CAGGSCCAGATGTGTGACCAAGGCAATCAGCCCGGCACTGTGAACGCTGTGGGTGAAGTGTG 492
Qy      466 ATCGAGAGAGAGGCGCTTACGCGCGGAGGTATCCCAATGTTCAACCGGCTGTGAGAGAGGCG 525
Db      493 GTGAGAGAGAGGCGCTTACGCGCGGAGGTATCCCAATGTTCAACCGGCTGTGAGAGAGGCG 552
Qy      526 GCCACCCCGCAGAGCTGTAACAGATGTTGAACAACCGTGGGCGGCAACAGAGCGCCATG 585
Db      553 GCCACCCCGCAGAGCTGTAACAGATGTTGAACAACCGTGGGCGGCAACAGAGCGCCATG 612
Qy      586 CAGATCTGAAGAGACACATCAACGAGAGGCGCGGAGTGTGAGACCGCTGTACCCGCTG 645
Db      613 CAGATCTGAAGAGAGACATCAACGAGAGGCGCGGAGTGTGAGACCGCTGTACCCGCTG 672
Qy      646 CAGGCGCGCGCGCGCGCGCGCGCAAGTGTGCGCAACCCCGCGGCAAGAGATATGCGCGG 705
Db      673 CAGGCGCGCGCGCGCGCGCGCGCAAGTGTGCGCAACCCCGCGGCAAGAGATATGCGCGG 732
Qy      706 GCCACCAAGACCTGTCAGAGAGAGATGCGCTGTGATGACCAAGAACCCCGCGTGGCGCTG 765
Db      733 ACCACCAAGACCTGTCAGAGAGAGATGCGCTGTGATGACCAAGAACCCCGCATATCCCGCTG 792
Qy      766 GCGGACATCTACAGCGGTGATCATCTGTGGCTGAAACAAGATGTGTGGATGTACAGC 825
Db      793 GCGGAGATCTACAGCGGTGATCATCTGTGGCTGAAACAAGATGTGTGGATGTACAGC 852
Qy      826 CCCGAGACATCTGAGAGATCCGCGCAGAGGCGCCCAAGAGGCGCTTCCGCGCACTAGCTGAG 885
Db      853 CCCACACATCTGAGAGATCCGCGCAGAGGCGCCCAAGAGGCGCTTCCGCGCACTAGCTGAG 912
Qy      886 CGCTTTCAAGACCTCTGCGCGGAGAGAGCCACCGAGAGTGAAGAACTGGATGACC 945
Db      913 CGCTTTCAAGACCTCTGCGCGGAGAGAGCCACCGAGAGTGAAGAACTGGATGACC 972
Qy      946 GAGACCTGTGTGTGAGAACCGCCACCCCGACTGTGACAACTCTGTGCGGCTCTCGGC 1005

```



Db 973 GAGACCTGTGTGTCAGAACGCCAACCCGACTGCAAGACATCTTGAAGGCTCTGGC 1032  
QY 1006 CCCGGGGCCACCTTGGAGAGATGATGACCGCTTGCAGGGCTGGGCGCCCGGCCAC 1065  
Db 1033 CCCGGGGCCACCTTGGAGAGATGATGACCGCTTGCAGGGCTGGGCGCCCGGCCAC 1092  
QY 1066 AAGGCCCGGTGTGGCCGAGGCGATGAGCAGG---CCAAACAGCGTGAACATCATGATG 1122  
Db 1093 AAGGCCCGGTGTGGCCGAGGCGATGAGCAGGCGAGACGACCCGGCGACATCATGATG 1152  
QY 1123 CAGAAAGCACTTCAAGGGCCCCCGCGCAAGTCAAGTGTCTCAACTGCGCAAGAG 1182  
Db 1153 CAGCGCGGCACTTCCGCAACAGCGGGAAGACCGTCAAGTGTCTCAACTGCGCAAGAG 1212  
QY 1183 GGGCCATGTGCCAAGAACTGCGCGCGCCCGCCGCAAGAGGCTGTGGAAGTGGCAAG 1242  
Db 1213 GGCACACCGCGCAGGAATGCGCGCGCCCGCCGCAAGAGGCTGTGCGCGCGCGCC 1272  
QY 1243 GAGGGCCACAGTGAAGACTGACCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGG 1302  
Db 1273 GAGGCAACCAATGAAGATGTCATGAGAGCAGGCTAATTTTAAAGAAATCTGG 1332  
QY 1303 CCCAGCCAAAGGCGCGCCCGCACTTCTGCAAGACCGGAGCCCGCGCGCC 1362  
Db 1333 CTTTCTTCAAGGGAAGGCGCAAGGAATTTTCTTCAAGCAG-----A 1374  
QY 1363 ACCGTGCCACCGCCCGCCCGCCGAGAGCTTCCGTTGAGAGACCAACCCCGCGCC 1422  
Db 1375 CCAGAGCCAAAGCCCGCCAGAGAGAGCTTCAAGTTGGGAGGAGAAACAATCTCC 1434  
QY 1423 AAGCAGAGCCCAAGAGACCGCGAGCGCTTACCGGAGCCCGCTGACCGCGCGAGCCG 1482  
Db 1435 TCTCAGAAACAGAGAGCCGATGACAGAACTGTATCTTTAACTTCCCTCAGATCATC 1494  
QY 1483 TTTCGAGCGCGCCCTCTGAGCCAGTAA 1509  
Db 1495 TTTGGCAAGACCCCTCTGTCAGAGTAA 1521

RESULT 9  
US-09-475-515-79  
Sequence 79, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 79  
LENGTH: 1865  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GP2  
US-09-475-515-79

Query Match 74.1%; Score 1117.8; DB 4; Length 1865;  
Best Local Similarity 85.1%; Pred. No. 1e-160;  
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

QY 1 ATGGGGCGCCCGCCGACGATCTCTGCGCGGAGAGAGCTGACAGAGTGGAGAAATCCGC 60

Db 13 ATGGGGCGCCCGCCGACGCTGTCAGCGCGGCGAGCTGACAAATGGAGAAATCCGC 72  
QY 61 CTGCGCCCGCGCGCAAGAGCACTAATGCTGAAGCACTGTGTGGCCAGCGCCAG 120  
Db 73 CTGCGCCCGCGCGCAAGAGCACTAATGCTGAAGCACTGTGTGGCCAGCGCCAG 132  
QY 121 CTGAGAGGCTTGGCTTGAACCCCGGCTGCTGAGACCGCGAGGCTGCAAGCATC 180  
Db 133 CTGAGAGGCTTGGCTTGAACCCCGGCTGCTGAGACCGCGAGGCTGCGCAGATC 192  
QY 181 ATGAAGAGCTGACCGCGCCCTGCAAGCCGCAACCGAGAGCTGCGAGCTGTAAC 240  
Db 193 CTGGCCAGCTGACCGCCAGCGCTGCAAGCCGCAAGAGCTGCGAGCTGTAAC 252  
QY 241 ACCGTGGCCACCTGTACTGCTGTGACCGCGCATGAGGTCTCGGACCAAGAGGCG 300  
Db 253 ACCGTGGCCACCTGTACTGCTGTGACCAAGGCAATGCAAGAGCAACAGAGAGGCG 312  
QY 301 CTGACAAAGATGAGAGAGAGCAAGCAATGCTCCAGCAAGAACCCAGAGCCAGAG 360  
Db 313 CTGAGAAAGATGAGAGAGAGCAAGCAATGCTCCAGCAAGAGCCAGAGGCGCCGCC 372  
QY 361 GCCGAGCGCA-----AGGTGAGCCAGAACTACCCCATGTGTGAGAACTCG 405  
Db 373 GCCGCGGCAACCGGCAAGACAGCGAGGTGAGCCAGAACTACCCCATGTGTGAGAACTCG 432  
QY 406 CAGGCGCAGATGTGTGACACAGGCGCATCAGCCCGCGCACTGAAAGCTGTGGTGAAGTG 465  
Db 433 CAGGCGCAGATGTGTGACACAGGCGCATCAGCCCGCGCACTGAAAGCTGTGGTGAAGTG 492  
QY 466 ATCGAGAGAGAGCTTCAAGCCCGAGGTGATCCCATGTTTCAACGCGCTTGAAGAGG 525  
Db 493 GTGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTTCAAGCGCTTGAAGAGG 552  
QY 526 GCCACCCCGCAAGCACTGAACAGATGTTGAACACCGGTGGCGGCGCACAGCGCGCC 585  
Db 553 GCCACCCCGCAAGCACTGAACAGATGTTGAACACCGGTGGCGGCGCACAGCGCGCC 612  
QY 586 CAGATGCTGAAGAGCAACCATCAACAGAGAGCGCGCGAGTGGAGCCGCTGACCCCGTG 645  
Db 613 CAGATGCTGAAGAGCAACCATCAACAGAGAGCGCGCGAGTGGAGCCGCTGACCCCGTG 672  
QY 646 CAGCGCGCGCGCGTGGCGCCCGCGCAAGTGGCGGACCCCGCGGAGCGACATGCGCG 705  
Db 673 CAGCGCGCGCGCGTGGCGCCCGCGCAAGTGGCGGAGCCCGCGGAGCGACATGCGCG 732  
QY 706 GCCACCAAGACCCGCGAGAGAGATGCGCTGAGTGAACAGCAACCCCGCGCGCGTG 765  
Db 733 ACCACCAAGACCCGCGAGAGAGATGCGCTGAGTGAACAGCAACCCCGCGCGCGTG 792  
QY 766 GCGCATCTTCAAGAGCGGTGATCATCTGGGCGCTGAACAAGATGTGCGATGTAACG 825  
Db 793 GCGGAGATCTAAGAGCGGTGATCATCTGGGCGCTGAACAAGATGTGCGATGTAACG 852  
QY 826 CCGGTAGCATCTGAGACATCGCGGCGCCCAAGAGGCGCTTCCGAGCTACGTGAGC 885  
Db 853 CCGACCAAGCATCTGAGACATCGCGGCGCCCAAGAGGCGCTTCCGAGCTACGTGAGC 912  
QY 886 CGCTTCTTCAAGACCCGCGAGAGAGCGCAAGCGCAAGCGTGAACAACTGATGAGC 945  
Db 913 CGCTTCTTCAAGACCCGCGAGAGAGCGCAAGCGCAAGCGTGAACAACTGATGAGC 972  
QY 946 GAGACCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005  
Db 973 GAGACCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032  
QY 1006 CCCGGCGCCACCTTGAAGAGATGATGACCGCTTGCAGGGCTGGGCGCCCGGCCAC 1065  
Db 1033 CCCGGCGCCACCTTGAAGAGATGATGACCGCTTGCAGGGCTGGGCGCCCGGCCAC 1092  
QY 1066 AAGGCCCGGTGTGGCCGAGGCGATGAGCAGG---CCAAACAGCGTGAACATCATGATG 1122

Db 1093 AAGCCCGCTGCTGCGCCAGAGCGATGAGCCAGTGAAGAACCCCGGACCATCATGATG 1152  
QY 1123 CAGAAAGCAATTCAAGGGGCCCCCGGCCCAAGTCAAGTGTCTTCAACTGGCGCAAGAG 1182  
Db 1153 CAGCGCGCAACTTCGCAACAGCCGAAAGACCTGAAGTGTCTTCAACTGGCGCAAGAG 1212  
QY 1183 GGCACATCGCCAGAACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAAGTGGCGCAAG 1242  
Db 1213 GGCACACCGCGAGAACTGCGCGCGCCCCCGCAAGAGGGCTGTGAGGCGTGGCGCGC 1272  
QY 1243 GAGGCGCCACATGAAGATGACGACCGAGCGCCAGGCCAATTCTGGGGCAAGTGG 1302  
Db 1273 GAAGGACACCAATGAATGAAATGTCATGAGAGACGGCTTAATTTTATAGGAAATCTGG 1332  
QY 1303 CCAGCCACAGAGGCGCGCGCCAGCACTTCTGCAAGAACCGAGCGAGCGCGCGCGCC 1362  
Db 1333 CTTCTTACAAAGGAAGGCGCAAGGAATTTTCTTCAAGACAG-----A 1374  
QY 1363 ACCGTGCCACCGCGCGCGCGCGCGAGACTTCCGCTTCAAGAGACACCGCGCGCGCC 1422  
Db 1375 CCAGAGCCAAAGCGCCACAGAAAGAGCTTCAAGTTTGGGGAGGAGAAACAATCTCC 1434  
QY 1423 AAGCAGAGGCCCAAGAGACCGCGAGCTTACCGGAGCGCTTACCGCGCTGGCGAGCTG 1482  
Db 1435 TCTCAAGAAAGCAGAGCGCGATGAGCAAGAACTGTATCTTTAACTTCCCTCAGATCATTC 1494  
QY 1483 TTGGGAGCGCGCGCGCTGAGCGATTA 1509  
Db 1495 TTGGCAAGCAGCCCTGTCAGAGTAA 1521

RESULT 10  
US-09-475-515-6  
Sequence 6, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 4319  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: HIV-Gag-polymerase  
US-09-475-515-6

Query Match 74.1%; Score 1117.8; DB 4; Length 4319;  
Best Local Similarity 85.1%; Pred. No. 1e-160;  
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

QY 1 ATGGGCGCGCGCGCGCATCTGCGCGCGAGAGAGCTTGACACAGTGGAGAAATCCGC 60  
Db 7 ATGGGCGCGCGCGCGCGCTGCTGAGCGGCGGCGAGACTGACAAAGTGGAGAAATCCGC 66  
QY 61 CTGGCGCGCGCGCGCGAGAAAGCACTACAGCTGAAGACACTGTGTGGCGCGAGCGCGAG 120  
Db 67 CTGGCGCGCGCGCGCGAGAAAGTACAAAGCTGAAGACACTGTGTGGCGCGAGCGCGAG 126

QY 121 CTGAGAGGCTTTCGCTGAAACCCCGCGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180  
Db 127 CTGAGAGGCTTTCGCTGAAACCCCGCGCTGTGAGAGACAGAGAGGCTGCGCAGATC 186  
QY 181 ATGAAGCACTGCAAGCGCGCGCTGCAAGCCGACACCGAGAGCTGCGCAGCTTGAAC 240  
Db 187 CTGGGCGAGCTGAGCGCCAGCTGACACCGCGAGCGGGGAGCTGCGCAGCTGTACAA 246  
QY 241 ACCGTGGCACCTGTATGCTGAGTGAAGCGCGCATTCGAGGTCGGGACACCAAGAGGCG 300  
Db 247 ACCGTGGCACCTGTATGCTGAGTGAAGCGCGCATTCGAGGTCGGGACACCAAGAGGCG 306  
QY 301 CTGACAAAGATCGAGAGAGCAGAAACAATGCCAGCAAGAAAGACCAAGAGGAG 360  
Db 307 CTGAGAAAGATCGAGAGAGCAGAAACAATGCCAGCAAGAAAGGCGCAGCGCGCGC 366  
QY 361 GCCGAGCGCA-----AGTGAAGCCAAATCTACCCCATGTGTGCAAACTTG 405  
Db 367 GCCGCGCGCACCGGACAGAGCGAGGTGAGCAAACTACCCCATGTGTGCAAACTTG 426  
QY 406 CAGGCGCAGATGTGACACAGGCAATCAGCCCGCGCACCTGAAGCGCTGGGTGAAGTG 465  
Db 427 CAGGCGCAGATGTGACACAGGCAATCAGCCCGCGCACCTGAAGCGCTGGGTGAAGTG 486  
QY 466 ATCGAGGAAGGCTTTCAGCGCGCGAGGTGATCCCATGTTCAACGCGCTTGAAGAGGCG 525  
Db 487 GTGAGAGGAAGGCTTTCAGCGCGCGAGGTGATCCCATGTTCAAGGCGCTTGAAGAGGCG 546  
QY 526 GCCACCGCGCGAGACTGAACACGATGTTGAACACCGTGGCGCGCACAGGCGCGCATG 585  
Db 547 GCCACCGCGCGAGACTGAACACGATGTTGAACACCGTGGCGCGCACAGGCGCGCATG 606  
QY 586 CAGATCTGAAGAGCACCATCAAGAGAGGCGCGGAGTGGAGCCGCTGACCCGCTG 645  
Db 607 CAGATCTGAAGAGCACCATCAAGAGAGGCGCGGAGTGGAGCCGCTGACCCGCTG 666  
QY 646 CAGGCGCGCGCGCGCGCGCGCGAGATGCGCAAGTCCCGCGCGAGCGACATCGCGCGC 705  
Db 667 CAGGCGCGCGCGCGCGCGCGCGAGATGCGCGCAAGTGGCGAGCCCGCGCGAGGACATCGCGCGC 726  
QY 706 GCCACAGACCGCTGCAAGAGACGATGCGCTGTGATGACAGCAACCCCGCTGCGCGTG 765  
Db 727 ACCACAGACCGCTGCAAGAGACGATGCGCTGTGATGACCAACCCCGCTGCGCGTG 786  
QY 766 GCGCATCTCAAGCGGTGATCATCTGCGCGCGCAAGATGCTGCGATGACAGC 825  
Db 787 GCGCATCTCAAGCGGTGATCATCTGCGCGCGCAAGATGCTGCGATGACAGC 846  
QY 826 CCGGTGACATCTGGAATCCGCGAGGCGCCCAAGAGGCGCTTCCGCGACTAGGTGAC 885  
Db 847 CCGACAGCATCTGGAATCCGCGAGGCGCCCAAGAGGCGCTTCCGCGACTAGGTGAC 906  
QY 886 CGCTTCTTCAAGACCTGCGCGCGCGAGGCGACCCAGAGCTGAAGAACTGGATGACC 945  
Db 907 CGCTTCTTCAAGACCTGCGCGCGCGAGGCGACCCAGAGCTGAAGAACTGGATGACC 966  
QY 946 GAGACCTGCTGAGTCAAGAGCGCAACCCCGAGCTGAAGACATCTGCGCGCTCGGC 1005  
Db 967 GAGACCTGCTGAGTCAAGAGCGCAACCCCGAGCTGAAGACATCTGAGAGCTCTCGGC 1026  
QY 1006 CCGGCGCGCACCTGAGAGATGATGACCGCTGCGCGAGGCGTGGCGCGCGCGCGCAC 1065  
Db 1027 CCGGCGCGCACCTGAGAGATGATGATGACCGCTGCGCGAGGCGTGGCGCGCGCGCAC 1086  
QY 1066 AAGGCGCGCTGCTGCGCGAGGCGAGTGAAGGCGCGCAAGCTGAAGTGAATG 1122  
Db 1087 AAGGCGCGCTGCTGCGCGAGGCGAGTGAAGGCGCGCAAGCTGAAGTGAATG 1146  
QY 1123 CAGAAAGCACTTCAAGAGGCGCGCGCGCAAGCTGAAGTGAATGAGTGAAGAGAG 1182  
Db 1147 CAGCGCGCAACTTCCGCAACAGGAGAAAGCGTCAAGTGTCTTCACTGGCGCGAGAG 1206  
QY 1183 GGCACATCGCCAGAACTGCGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGGCGCAAG 1242

```

Db      1207 GGGCAACCGCCAGGAACCTGCGCGCGCCCGCCCGCAAGAGGCTGCTGGCGCTGGCCCGC 1266
QY      1243 GAGGCGCAACGATGAGAGACTGACCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGG 1302
Db      1267 GAAAGGACCAATGAAATGATGCTAGAGAGCAAGGCTTAATTTTAAAGGAAATCTGG 1326
QY      1303 CCCAGCAAGAGGCGCGCCCGCCCACTTCTGGCAAGCCAGCCAGCGCCCGCGCCCGC 1362
Db      1327 CTTCTTACAGAGGAAGGCGAGGAAATTTTCTTCAAGAGAG-----A 1368
QY      1363 ACCGTGCCCCAGCGCCCCCGCCGAGAGCTTCCGTTGAGAGAGCAACCCCGCGCCCGC 1422
Db      1369 CCAGAGCCAAACAGCCCCACCAAGAAAGAGCTTCAAGTTTGGGAGAGAGAAAACAATCTCC 1428
QY      1423 AAGCAGAGAGCCCAAGAGACCGCGAGCCCTTACCGGAGCCCTTGACCGCTTGGCAGCTTG 1482
Db      1429 TCTCAGAAACAGAGAGCCGATAGCAAGAACTGTATCTTTAACTTCCCTCAGATCAGTTC 1488
QY      1483 TTCCGAGGCGCGCCCTGAGCCAGTAA 1509
Db      1489 TTGGCAACGACCCCTGTCAAGTAA 1515

RESULT 11
US-09-475-515-9
/ Sequence 9, Application US/09475515A
/ Patent No. 6602705
/ GENERAL INFORMATION:
/ APPLICANT: BARNETT, Susan
/ APPLICANT: ZUR MEGEDE, Jan
/ APPLICANT: SRIVASTAVA, Indresh
/ APPLICANT: LIAN, Ying
/ APPLICANT: HARTOG, Karin
/ APPLICANT: LIU, Hong
/ APPLICANT: GREER, Catherine
/ APPLICANT: SELBY, Mark
/ APPLICANT: WALKER, Christopher
/ TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
/ FILE REFERENCE: 1621.002
/ CURRENT APPLICATION NUMBER: US/09/475,515A
/ CURRENT FILING DATE: 1999-12-30
/ NUMBER OF SEQ ID NOS: 90
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 9
/ LENGTH: 1268
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
/ OTHER INFORMATION: common region
US-09-475-515-9

Query Match      69.5%; Score 1049.2; DB 4; Length 1268;
Best Local Similarity 90.8%; Pred. No. 2.4e-150;
Matches 1146; Conservative 0; Mismatches 98; Indels 18; Gaps 2;

QY      1  ATGGGCGCCCGCGCGCATCTGCGCGCGAGAGCTGAGCAAGTGGGAGAAATCCGC 60
Db      7  ATGGGCGCCCGCGCGCATCTGCGCGCGAGCGCGAGCTGAGCAAGTGGGAGAAATCCGC 66
QY      61  CTGGCGCCCGCGCGCAAGAGCACTTACATGCTGAAGCACCTGTGTGTGGCCAGCCGCGAG 120
Db      67  CTGGCGCCCGCGCGCAAGAGTACAAAGCTGAACATCGTGTGTGGCCAGCCGCGAG 126
QY      121  CTGGAAGGCTTGGCCCTGTAACCCCGGCTGTGGAACCCCGAGAGGCTGCAAGCAGATC 180
Db      127  CTGGAAGGCTTGGCCCTGTAACCCCGGCTGTGGAACCCCGAGAGGCTGCGCCAGATC 186
QY      181  ATGAAGCAGCTGAGCCCGCCTGCAAGACCGGACCGAGAGAGCTGGCGAGCCTGTACAAC 240
Db      187  CTGGGCACTGAGCCCGCAGCCTGCAAGACCGGAGGAGAGCTGGCGAGCCTGTACAAC 246
```

```

QY      241  ACCGTGCGCAACCTGTATCTGCTGCAACCGCGCATTCGAGTCCGCGACCAAGAGGCTC 300
Db      247  ACCGTGCGCAACCTGTATCTGCTGCAACCGCGCATTCGAGTCCGCGACCAAGAGAGGCTC 306
QY      301  CTGGAAGATGAGAGAGAGAGAGCAAGTCCAGCAGAAAGACCAGAGGCGCAAGAG 360
Db      307  CTGGAAGATGAGAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY      361  GCGGACGGCA-----AGTGAACCAAGTACCCATGCTGAGAACTTG 405
Db      367  GCGGCGGAGACCGGCAACAGCAGCAGAGGTAGACCAAGACTACCCATGCTGAGAACTTG 426
QY      406  CAGGCGCAGATGATGATCAACAGGCGCATCAGCCCCCGCACTTGAAGGCTTGGTGAAGTG 465
Db      427  CAGGCGCAGATGATGATCAACAGGCGCATCAGCCCCCGCACTTGAAGGCTTGGTGAAGTG 486
QY      466  ATCGAGAGAAAGGCTTCAAGCCCGAGGATGCCCATGTTTCAACCGGCTTGAAGAGGAGC 525
Db      487  GTGAGAGAGAGAGGCTTCAAGCCCGAGGATGCCCATGTTTCAAGGCGCTTGAAGAGGAGC 546
QY      526  GCGACCCCGCAGAGACCTGAAACAAGTGTGAACAAGTGGGCGGCAACAGGCGCGCATG 585
Db      547  GCGACCCCGCAGAGACCTGAAACAAGTGTGAACAAGTGGGCGGCAACAGGCGCGCATG 606
QY      586  CAGATGCTGAAGAGACCATCAACGAGAGAGCGCCGCGAGTGGAGACCGCTTGACCCCTTG 645
Db      607  CAGATGCTGAAGAGACCATCAACGAGAGAGCGCCGCGAGTGGAGACCGCGCTTGACCCCTTG 666
QY      646  CAGGCGGCGCGCGGCGCGCGCGAGTGGCGGAGCCCGCGGAGAGGACATCGCGCGGCG 705
Db      667  CAGGCGGCGCGCGCGCGCGCGCGAGTGGCGGAGCCCGCGGAGAGGACATCGCGCGGCG 726
QY      706  GCGACAGAGACCCCTGAGAGAGAGATCGCTGTGATGACCAAGCAACCCCGCTGCGCGTG 765
Db      727  ACCACAGAGACCCCTGAGAGAGAGATCGCTGTGATGACCAACCAACCCCGCTGCGCGTG 786
QY      766  GCGGACATTCAGAGCGGTGATCATCTGGGCGCTGAAACAGATGTGTGGATGTACAGC 825
Db      787  GCGGAGATTCAGAGCGGTGATCATCTGGGCGCTGAAACAGATGTGTGGATGTACAGC 846
QY      826  CCGGAGAGATCTGAGACATCGCGCAAGGCGCCCAAGAGCCCTTCCGAGCATGAGTGCAC 885
Db      847  CCGACAGAGATCTGAGACATCGCGCAAGGCGCCCAAGAGCCCTTCCGAGCATGAGTGCAC 906
QY      886  CGCTTCTTCAAGAGCCCTGCGCGCGAGAGGCGCAACCGAGAGCTGAAGAACTGATGACC 945
Db      907  CGCTTCTTCAAGAGCCCTGCGCGCGTGAAGAGGCGCAACCGAGAGCTGAAGAACTGATGACC 966
QY      946  GAGACCTGTGTGTGAGAGAGCGCAACCCCGAGCTGAAGAGCAATCTTGGCGCTCTCGGC 1005
Db      967  GAGACCTGTGTGTGAGAGAGCGCAACCCCGAGCTGAAGAGCAATCTTGAAGGCTCTCGGC 1026
QY      1006  CCGGCGGCAACCTGAGAGAGTGAATGACCGGCTGCGAGGCGTGGGCGCGCCCGCGCAC 1065
Db      1027  CCGGCGGCAACCTGAGAGAGTGAATGACCGGCTGCGAGGCGTGGGCGCGCCCGCGCAC 1086
QY      1066  AAGGCGCGGCTGTGGCGGAGGCGAGTGAAGCGAG---CCAAGCGGTGAACATCTGATG 1122
Db      1087  AAGGCGCGGCTGTGGCGGAGGCGAGTGAAGCGAGCGGCGAGCAATCTGATGATG 1146
QY      1123  CAGAAAGCAATTCAGAGGCGCCCGGCGCAAGTCAAGTGTTCACATGCGCGCAAGAG 1182
Db      1147  CAGCGCGGCAATTCAGCAACAGAGGAGAAAGACCGTCAAGTGTTCACATGCGCGCAAGAG 1206
QY      1183  GCGCAATTCGCAAGAACTGCGCGCGCCCGCGCAAGAGGCGTGTGGAAGTGGCGCAAG 1242
Db      1207  GCGCAATTCGCAAGAACTGCGCGCGCCCGCGCAAGAGGCGTGTGCGCGCTGCGCGCGC 1266
QY      1243  GA 1244
Db      1267  GA 1268
```



```

RESULT 13
US-09-974-702-1
; Sequence 1, Application US/09974702
; Patent No. 6696291
; GENERAL INFORMATION:
; APPLICANT: Shiver, John W.
; Davies, Mary Ellen
; Freed, Daniel C.
; Liu, Margaret A.
; Perry, Helen C.
; TITLE OF INVENTION: Synthetic HIV Gag Genes
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/974, 702
; FILING DATE: 09-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,981
; FILING DATE: 03-FEB-1998
; APPLICATION NUMBER: US60/037,854
; FILING DATE: 07-FEB-1997
; APPLICATION NUMBER: GB9705040.5
; FILING DATE: 12-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19730
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; TEXT: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-974-702-1

Query Match          64.4%; Score 971.4; DB 4; Length 1532;
Best Local Similarity 79.6%; Pred. No. 14e-138;
Matches 1210; Conservative 0; Mismatches 281; Indels 30; Gaps 4

QY      1 ATGGGCGCCGCCAGCATCTTGCGCGCGCGAAGCTGGAACAATGTGAGAATTCGC 60
DB      10 ATGGTGTCTAAGGCTTTCTGTCTGTCTGTGTGTGAGCTGGAACAATGTGAGAATTCAG 69
QY      61 CTGCGCCCCGCGCGCAAGAACAATCATCTGAAACACTTGTTGTGGCCAGCCGCGAG 120
DB      70 CTGAAGCCTGTGGCGCAAGAAGAATCAAGCTTAAGCAATTGTGTGGGCTTCCAGGGAG 129
QY      121 CTGAGAGGCTTGGCCCTTAGAACCCCGGCTGTGAGAACCGCGAGGGGCTTGAAGAGATC 180
DB      130 CTGAGAGAGTTTGTCTGTGAACCTTGCCCTGTGCTGTGAGAACCTTCTGAGGGGCTGCAAGGAGATC 189
QY      181 ATGAAGCAGCTGACAGCCGCGCTTGAGAACCGGCACTCGAAGAGCTGTGGCAGCTGTAAAC 240

```

Db	130	CTGGGCGCAGCTTCCAGCCTCTCCCTGGAAACAGGCTCTGAGGAGCTAAGGTCCCTGTACAC	249
Qy	241	ACCGTGGCCACCCTGTACTGCGGTGCAGCGCCGGCATCGAGGTCCGGGACACCAAGAGGCC	300
Db	250	ACAGTGGCTACCTGTACTGTGTGTGCACCAAGAAATTGATGTGAAGGACACCAAGAGGCC	309
Qy	301	CTGGACAAAGATCGAGAGAGAGCAAAAGTCCAGCGAGAGACCACAGCAGCC-----	353
Db	310	CTGGAGAAAGATTGAGAGAGAGAGCAAAAGTCCAAAGAAAGGCCACGACGCTGTGCT	369
Qy	354	--CAAGAGGGCCGACGGGAGGTGAGCCAGAACTACCCCATGTGTCAAGAACTCGAGGGC	411
Db	370	GGCACAAGCAACTCCAGCGAGGTGTCCAGAACTACCCCATGTGTCAAGAACTCCAGGGCC	429
Qy	412	CAGATGTGTACACAGGACCATCAGCCCCCGACCCCTGAACCGCCTGGGTGAAGGTGATCGAG	471
Db	430	CAGATGTGTACACAGGACCATCTCCCCCGAGCCCTGAATGCTCGGTGAAGGTGTGAGAG	489
Qy	472	GAGAAAGCCTTTCAGGCCCGGAGGTATCCCATGTTTCAACCGCCTTGAGCGAGGGCCGAC	531
Db	490	GAGAAAGCCTTCTCTCCCTGAGGTGATCCCATGTTTCTCTGCCCTGTGAGGGTCCAC	549
Qy	532	CCCCGAGACCTGGAACACGATGTTGAAACACGTGGGCGGCCACAGGCCGCATATGACATG	591
Db	550	CCCCGAGACCTGGAACACGATCTGAAACAGGTGGGGGCGCATACAGGCTGCCATATGACATG	609
Qy	592	CTGAAGGACACCATCATCAAGAGAGGCGCCCGAGTGGGACCGCCTGTCAACCCGCTGACGCC	651
Db	610	CTGAAGGAGACCATCATATGAGAGAGGCTGTGATGTGGACAGGCTGTGATCTGTGACAGCT	669
Qy	652	GGCCCCGTGGCCCCCGGCGCAGATGCGACACCCCGCGGCAAGGACATCGCCGCGCCAC	711
Db	670	GGCCCCCATTTGCCCGCGGCAGATGAGGAGCCCAAGGGCTGTGACATGTGTGGACACAC	729
Qy	712	AGCACCTCTGCAGAGACAGATGCGCTGTGATGACACGACACCCCTCCGTGCGGTGGCCGAC	771
Db	730	TCCACCCCTCCAGAGACAGATTTGGCTGATGACCAACACCCCCCATCCTGTGTGGGGGA	789
Qy	772	ATCTCAAGCGGTGATCATCTCGGGGCTTGAAACAAGATCGTGGCGATGTACAGCCCGGTG	831
Db	790	ATCTCAAGAGGTGATCATCTGTGGGCTTGAAACAAGATTTGTAGGATGTATCTCCCCAC	849
Qy	832	AGCATCTCTGACATCTCGCCAGGGCCCCCAAGAGGCCCTTCCGCACTACGTGACCGCTTC	891
Db	850	TCCATCTCTGACATATAGGCGAGGGCCCCCAAGAGGCCCTTCAAGGACTATGTGGACAGTTTC	909
Qy	892	TTCAAGACCTCTGCGCGCGCGAGCGAGGCCAACCAAGACGTGAAGAACTGATGACCGAGACC	951
Db	910	TACAAGACCTCTGAGGGCTGACAGGCCCTCCAGAGGGTGAAGAACTGATGACAGAGACC	969
Qy	952	CTGCTGTGTGCAAGAAAGCCCAACCCCGACCTGCAAGAACATCTCTGCGGCTCTCGGCCCGGC	1011
Db	970	CTGCTGTGTGCAAGAAAGTCCCAACCTGTGACTGCAAGAACATCTCTGAAAGGCCCTGTGGCCCTCT	1023
Qy	1012	GCCACCTCTGAGAGAGATGATGACCGCTCGCCAGAGGCGTGGGCGGCCCGGCCCAAGAGGCC	1071
Db	1030	GCCACCTCTGAGAGAGATGATGACCGCTCGCCAGAGGCGTGGGCGGCCCTGTGTCAAGAGGCC	1089
Qy	1072	GCGGTGTGGCCGAGGCGATGAGCCAGG--CAAACAGCGTGAACATCATGATGTGCAAG	1128
Db	1090	AGGGGTGTGGCTGAGGCGCATGTGCCAGGTGACCAATCTCGCCACATCATATGTGCAAGG	1144
Qy	1129	AGCAACTTCAGAGGCCCCCGGCGCAAGCTCAAGTCTTCAACTGCGGCAAGAGAGGCCAC	1188
Db	1150	GGCAACTTCAGGAACCAAGAGAGCAATGTAAGTCTTCAACTGTGGGCAAGGTGGGCCAC	1203
Qy	1189	ATGCGCAAGACTGACCGGCGCCCCCGCAAGAGGGCTGTGGAAGTGTGGCAAGAGAGGCC	1248
Db	1210	ATTGCGCAAGACTGTAGGGCCCCCAAGAGAGAGGGCTGTGGAAGTGTGGCAAGAGAGGCC	1266
Qy	1249	CACCAAGATGAAGACTGACCCGAGCGCCAGGCCCACTTCTGTGGCAAGATCTGGCCGAC	1308



Oy 1389 CCCACGGGCCCCCGCGGAGGCTTGGCTTGCAGAGAGACACACCCCGCCCAAGCAG 1428  
 Db 1390 GAGTCTTTC-----AGTTTGGGAGAGGAAGACACCCCGACAGAAAGCAG 1437  
 Oy 1429 GAGCCCAAGGACCGCGAGCCCTTACCGCGAGGCCCTGCGCGAGGCTGTGGC 1488  
 Db 1438 GAGCCCATTTGCAGAGAGAGCTGTAC-----CCCTCGGCTCTCTCGAGGTCCTGTGTGGC 1491  
 Oy 1489 AGCGGCCCTGAGCCAGTAA 1509  
 Db 1492 AACGACCTTCTCTCCAGTAA 1512

RESULT 15  
US-09-952-060-25

Query Match	64.4%	Score	971.4	DB	4	Length	37474
Best Local Similarity	79.6%	Pred. NO.	1.3e-138				
Matches 1210; Conservative	0	Mismatches	281			Indels	30; Gaps 4

[illegible]

```

Qy      1369  CCCACCGCCCCCCCCCGGAGAGCTTCCGCTTCGAGAGAGACCACCCCGCGCCCCCAAGCAG 1428
Db      2652  GAGTCCTTC-----AGGTTGGGAGAGAGAGACCACCCCGAGCCAGAGAGCAG 2699

Qy      1429  GAGCCCAAGAGACCGGAGCCCTACCGGAGACCCCTGAGCCGCTGCGCAGCCTGTTCGGC 1488
Db      2700  GAGCCCATTGACAGAGAGCTGTAC-----CCCTGGCCTCCCTGAGGTCCCTGTTTGGC 2753

Qy      1489  AGCGGCCCCCCTGAGCCAGTAA 1509
Db      2754  AACGACCCCTCCTCCCACTAA 2774
  
```

Search completed: September 1, 2005, 22:46:26  
 Job time : 274.651 secs



1	-1509	100.0	1509	10	US-09-899-576-64	Sequence 4, Apr1
2	1507.4	99.9	1509	10	US-09-967-466-64	Sequence 64, Apr1
3	1504.2	99.7	1509	10	US-09-967-466-68	Sequence 68, Apr1
4	1489.8	98.7	1509	10	US-09-899-576-21	Sequence 21, Apr1
5	1297.6	85.0	1494	10	US-09-899-576-51	Sequence 51, Apr1
6	1294.8	85.0	1491	10	US-09-899-576-99	Sequence 99, Apr1
7	1276.8	84.6	1497	10	US-09-967-466-63	Sequence 63, Apr1

8	1276.8	84.6	3167	10	US-09-899-575-3	Sequence 3,	Appl1
9	1275.2	84.5	3162	15	US-10-190-435-18	Sequence 18,	Appl1
10	1275.2	84.5	3462	15	US-10-190-435-16	Sequence 16,	Appl1
11	1275.2	84.5	4419	15	US-10-190-435-19	Sequence 19,	Appl1
12	1275.2	84.5	4419	16	US-10-190-3058-14	Sequence 14,	Appl1
13	1275.2	84.5	4683	15	US-10-190-435-35	Sequence 35,	Appl1
14	1275.2	84.5	4606	15	US-10-190-435-34	Sequence 34,	Appl1
15	1275.2	84.5	4615	15	US-10-190-435-36	Sequence 36,	Appl1
16	1275.2	84.5	4702	15	US-10-190-435-38	Sequence 38,	Appl1
17	1275.2	84.5	4716	15	US-10-190-435-17	Sequence 17,	Appl1
18	1275.2	84.5	4716	16	US-10-190-3058-13	Sequence 13,	Appl1
19	1273.6	84.4	2742	15	US-10-190-435-20	Sequence 20,	Appl1
20	1273.6	84.4	2742	16	US-10-190-3058-15	Sequence 15,	Appl1
21	1273.6	84.4	3930	15	US-10-190-435-9	Sequence 9,	Appl1
22	1273.6	84.4	3930	15	US-10-190-435-10	Sequence 10,	Appl1
23	1273.6	84.4	3930	15	US-10-190-435-11	Sequence 11,	Appl1
24	1273.6	84.4	5145	15	US-10-190-435-12	Sequence 12,	Appl1
25	1273.6	84.4	5145	16	US-10-190-3058-12	Sequence 12,	Appl1
26	1273.6	84.4	4713	15	US-10-190-435-59	Sequence 59,	Appl1
27	1272.6	84.3	5184	15	US-10-190-3058-83	Sequence 83,	Appl1
28	1271.6	84.3	5184	16	US-10-190-435-58	Sequence 58,	Appl1
29	1271.6	84.3	5184	16	US-10-190-3058-82	Sequence 82,	Appl1
30	1271.2	84.2	1479	10	US-09-367-464-67	Sequence 67,	Appl1
31	1270.2	84.2	1742	15	US-10-190-435-57	Sequence 57,	Appl1
32	1270.2	84.2	1742	16	US-10-190-3058-81	Sequence 81,	Appl1
33	1260.8	83.6	1479	10	US-09-899-575-20	Sequence 20,	Appl1
34	1201	79.6	1515	17	US-10-387-336-4	Sequence 4,	Appl1
35	1201	79.6	4472	17	US-10-387-336-75	Sequence 75,	Appl1
36	1201	79.6	4608	17	US-10-387-336-76	Sequence 76,	Appl1
37	1201	79.6	4689	17	US-10-387-336-73	Sequence 73,	Appl1
38	1201	79.6	4766	17	US-10-387-336-74	Sequence 74,	Appl1
39	1199.4	79.5	2799	16	US-10-241-009-18	Sequence 18,	Appl1
40	1199.4	79.5	2799	16	US-10-190-434B-18	Sequence 18,	Appl1
41	1199.4	79.5	3205	16	US-10-190-3058-11	Sequence 11,	Appl1
42	1199.4	79.5	3205	16	US-10-241-009-17	Sequence 17,	Appl1
43	1199.4	79.5	3205	16	US-10-190-434B-17	Sequence 17,	Appl1
44	1199.4	79.5	3496	16	US-10-241-009-15	Sequence 15,	Appl1
45	1199.4	79.5	3496	16	US-10-190-434B-15	Sequence 15,	Appl1

## RESULT 1

```

US-09-899-575-4
: Sequence 4, Application US/09899575
: Publication No. US20030223961A1
:
GENERAL INFORMATION:
:
APPLICANT: zur Megede, Jan
APPLICANT: Barnett, Susan W.
APPLICANT: Egnelbrecht, Susan
APPLICANT: van Rensburg, Estrelita Janse
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: P01631.102
CURRENT APPLICATION NUMBER: US/09/899,575
CURRENT FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 09/475,704
PRIOR FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 135
SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 4
:
: LENGTH: 1509
:
: TYPE: DNA
:
: ORGANISM: Artificial Sequence
:
FEATURES:
:
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
OTHER INFORMATION: of HIV strain AF110967
US-09-899-575-4

Query Match          100.0%; Score 1509; DB 10; Length 1509;
Best Local Similarity 100.0%; Pctd. No. 0;
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

```

QY 1 ATGGGCGCCGCGCCGAGCATCTCTGGCGGGGAGAACTGGAGCAAGTGGGAGAAAGATCCGC 60
DB 1 ATGGGCGCCGCGCCGAGCATCTCTGGCGGGGAGAACTGGAGCAAGTGGGAGAAAGATCCGC 60
QY 61 CTGGCGCCCGGCGGCAAGAGCACTACATGCTGAAACACCTGGTGTGGCCAGCCCGGAG 120
DB 61 CTGGCGCCCGGCGGCAAGAGCACTACATGCTGAAACACCTGGTGTGGCCAGCCCGGAG 120
QY 121 CTGGAGGGCTTGGCCCTGAAACCCCGGCTGCTGGAGAACCCCGAGGGGCTGCAAGCAGATC 180
DB 121 CTGGAGGGCTTGGCCCTGAAACCCCGGCTGCTGGAGAACCCCGAGGGGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGCTGAGCCGCCCTGAGACCGGACCGAGAGAGCTGGCGAGCCCTGTACAC 240
DB 181 ATGAAGCAGCTGAGCCGCCCTGAGACCGGACCGAGAGAGCTGGCGAGCCCTGTACAC 240
QY 241 ACCGTGGCCACCTGTACTGCGTGCAAGCGCGGATGAGGTCCGCGACACCAAGAGGCC 300
DB 241 ACCGTGGCCACCTGTACTGCGTGCAAGCGCGGATGAGGTCCGCGACACCAAGAGGCC 300
QY 301 CTGGACAAATGAGAGAGAGAGCAAGAACTCCAGACAGAAACCCAGAGGCCAAGAG 360
DB 301 CTGGACAAATGAGAGAGAGAGCAAGAACTCCAGACAGAAACCCAGAGGCCAAGAG 360
QY 361 GCCGACGGGAGAGTGAGCCAGAACTACCCCATGTGCAAGAACTTGAGGGGCCAGATGATG 420
DB 361 GCCGACGGGAGAGTGAGCCAGAACTACCCCATGTGCAAGAACTTGAGGGGCCAGATGATG 420
QY 421 CACGAGGCGATAGCCCGCCGACCTGAAAGCGCTGGGTGAGGTGATCGAGAGAGAGCC 480
DB 421 CACGAGGCGATAGCCCGCCGACCTGAAAGCGCTGGGTGAGGTGATCGAGAGAGAGCC 480
QY 481 TTCAAGCCCGGAGGTGATCCCATGTTCATCCGCTTACAGGAGAGGGGCCACCCCGCAGAG 540
DB 481 TTCAAGCCCGGAGGTGATCCCATGTTCATCCGCTTACAGGAGAGGGGCCACCCCGCAGAG 540
QY 541 CTGAACACGATGTGAACACCGTGGGCGGCGACACAGGCGCCATGCGAGATGCTGAAGAGAC 600
DB 541 CTGAACACGATGTGAACACCGTGGGCGGCGACACAGGCGCCATGCGAGATGCTGAAGAGAC 600
QY 601 ACCATCAACGAGAGAGCGCGGAGTGGAGACCGCTTGACACCCCGTGAGGCGCGGCCCTG 660
DB 601 ACCATCAACGAGAGAGCGCGGAGTGGAGACCGCTTGACACCCCGTGAGGCGCGGCCCTG 660
QY 661 GCGCCCGGCGAGATGGGCGACCCCGCGGACAGAGATGCGCGGCGCCACACGACACCTG 720
DB 661 GCGCCCGGCGAGATGGGCGACCCCGCGGACAGAGATGCGCGGCGCCACACGACACCTG 720
QY 721 CAGAGACGATGCGCTGATGAGACAGACACCCCGCTGCGGTGGGCGAGATCTTACAG 780
DB 721 CAGAGACGATGCGCTGATGAGACAGACACCCCGCTGCGGTGGGCGAGATCTTACAG 780
QY 781 CGGTGATCATCTGTGGCTGAAACAGATGCTGCGGATGACAGCCCGTGAGCATCTG 840
DB 781 CGGTGATCATCTGTGGCTGAAACAGATGCTGCGGATGACAGCCCGTGAGCATCTG 840
QY 841 GACATCCGCGCAGAGGCGCCCAAGAGACCTTTCGCGACTAGTGAACCGCTTTCAGAGAC 900
DB 841 GACATCCGCGCAGAGGCGCCCAAGAGACCTTTCGCGACTAGTGAACCGCTTTCAGAGAC 900
QY 901 CTGGCGCGCGAGACGAGCCAGACGAGAGCTGAAGAACTGATGACGAGACCTTGTGCTG 960
DB 901 CTGGCGCGCGAGACGAGCCAGACGAGAGCTGAAGAACTGATGACGAGACCTTGTGCTG 960
QY 961 CAGAAAGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTGCGGCCCGGCGACACCTG 1020
DB 961 CAGAAAGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTGCGGCCCGGCGACACCTG 1020
QY 1021 GAGGAGATGATGACCGGCTGCGGAGGCGGTGGGCGGCCCGGCGACCAAGGCGCGCTGAG 1080
DB 1021 GAGGAGATGATGACCGGCTGCGGAGGCGGTGGGCGGCCCGGCGACCAAGGCGCGCTGAG 1080

```

```

QY 1081 GCCGAGCGATGAGGCCAGGCCCAAGCGGTGAACATCATGATGACAGAGCAACTTCAAG 1140
DB 1081 GCCGAGCGATGAGGCCAGGCCCAAGCGGTGAACATCATGATGACAGAGCAACTTCAAG 1140
QY 1141 GCGCCCGGCGCAAGCTCAAGTCTTCAATGCGGCAAGAGGGCCACATGCGCAAGAAC 1200
DB 1141 GCGCCCGGCGCAAGCTCAAGTCTTCAATGCGGCAAGAGGGCCACATGCGCAAGAAC 1200
QY 1201 TGGCGGCGCCCGGCAAGAGGCTGTGGAAGTGGCGGCAAGAGGGCCACAGATGAAG 1260
DB 1201 TGGCGGCGCCCGGCAAGAGGCTGTGGAAGTGGCGGCAAGAGGGCCACAGATGAAG 1260
QY 1261 GACTGACCGAGCGGCAAGCCCACTTCTGGGCAAGATCTGAGCCAGCCACAGAGGCCG 1320
DB 1261 GACTGACCGAGCGGCAAGCCCACTTCTGGGCAAGATCTGAGCCAGCCACAGAGGCCG 1320
QY 1321 CCGGCAACTTCTGCAAGAACCGAGAGGCCGCGCCGCCACCGTCCACCGCCGCC 1380
DB 1321 CCGGCAACTTCTGCAAGAACCGAGAGGCCGCGCCGCCACCGTCCACCGCCGCC 1380
QY 1381 CCGGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGGCCCAAGCAGAGGCCAAGAGAC 1440
DB 1381 CCGGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGGCCCAAGCAGAGGCCAAGAGAC 1440
QY 1441 CGGAGCCCTACCGGAGCCCTGACCGCCCTGCGAGACCTGTTGCGAGCGGCCCTG 1500
DB 1441 CGGAGCCCTACCGGAGCCCTGACCGCCCTGCGAGACCTGTTGCGAGCGGCCCTG 1500
QY 1501 AGCCAGTAA 1509
DB 1501 AGCCAGTAA 1509

```

```

RESULT 2
US-09-967-464-64
; Sequence 64, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: O'Brien, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Umer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PPI6269.004
; CURRENT FILING DATE: 2002-04-11
; PRIOR FILING DATE: 2000-09-28
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-64

```

```

Query Match 99.9%; Score 1507.4; DB 10; Length 1509;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGCCGCGCGAGCATCTGCGCGGCGAGAGAGTGAACAAGTGGAGAAAGATCCGC 60
DB 1 ATGGGCGCCGCGCGAGCATCTGCGCGGCGAGAGAGTGAACAAGTGGAGAAAGATCCGC 60
QY 61 CTGGCGCCCGGCGGCAAGAGCACTACATGCTGAAACACCTGGTGTGGCCAGCCCGGAG 120
DB 61 CTGGCGCCCGGCGGCAAGAGCACTACATGCTGAAACACCTGGTGTGGCCAGCCCGGAG 120

```

QY 121 CTGAGAGGCTTGCCCTGAAACCCCGGCTGCTGAGAGCCGCGAGGGCTGCAAGCAGATC 180  
 DB 121 CTGAGAGGCTTGCCCTGAAACCCCGGCTGCTGAGAGCCGCGAGGGCTGCAAGCAGATC 180  
 QY 181 ATGAACAGCTGAGCCCGCTGCAAGCCGAGCCAGAGAGCTGCGCAGCCTGTACAAC 240  
 DB 181 ATGAACAGCTGAGCCCGCTGCAAGCCGAGCCAGAGAGCTGCGCAGCCTGTACAAC 240  
 QY 241 ACCGTGGCCACCTGTACTGCTGCAAGCCGCGCATGAGAGTCCGCGACCAAGAGGCC 300  
 DB 241 ACCGTGGCCACCTGTACTGCTGCAAGCCGCGCATGAGAGTCCGCGACCAAGAGGCC 300  
 QY 301 CTGGAACAATGAGGAGGAGAGCAAGTCCACAGAAAGCCAGAGCCAGAGAG 360  
 DB 301 CTGGAACAATGAGGAGGAGAGCAAGTCCACAGAAAGCCAGAGCCAGAGAG 360  
 QY 361 GCCGACGGAGAGTGAAGCAGAACTACCCATCTGTCAGAACTGCGAGGGCCAGATGTG 420  
 DB 361 GCCGACGGAGAGTGAAGCAGAACTACCCATCTGTCAGAACTGCGAGGGCCAGATGTG 420  
 QY 421 CACGAGGCCATAGCCCCCGCACCTGAAAGCCTGAGTGAAGTATCAAGAAAGGCC 480  
 DB 421 CACGAGGCCATAGCCCCCGCACCTGAAAGCCTGAGTGAAGTATCAAGAAAGGCC 480  
 QY 481 TTACAGCCCGAGAGTGAATCCCATGTTCAACCGCCCTGAGAGAGGGCCACCCCGAGAC 540  
 DB 481 TTACAGCCCGAGAGTGAATCCCATGTTCAACCGCCCTGAGAGAGGGCCACCCCGAGAC 540  
 QY 541 CTGAACACAGATGTTGAACAACCGTGGGCGGCACAGCCGCCATGAGATGCTGAAGAAC 600  
 DB 541 CTGAACACAGATGTTGAACAACCGTGGGCGGCACAGCCGCCATGAGATGCTGAAGAAC 600  
 QY 601 ACCATCAACGAGAGAGCCCGCCAGTGGACCCGCTGACACCCGTCAGAGCCGCGCCGTG 660  
 DB 601 ACCATCAACGAGAGAGCCCGCCAGTGGACCCGCTGACACCCGTCAGAGCCGCGCCGTG 660  
 QY 661 GCGCCCGGAGATGGCGCAACCCCGCGGAGAGATATGCGGGCCACAGCAACCTTG 720  
 DB 661 GCGCCCGGAGATGGCGCAACCCCGCGGAGAGATATGCGGGCCACAGCAACCTTG 720  
 QY 721 CAGAGCAGATGCGCTGATGAGCAGCAACCCCGCGTCCGCGAGGAGCATCTACAAG 780  
 DB 721 CAGAGCAGATGCGCTGATGAGCAGCAACCCCGCGTCCGCGAGGAGCATCTACAAG 780  
 QY 781 CGGTGATCATCTGGGCTGAAACAAGATCGTGGATGACAGCCCGTGAAGATCTTG 840  
 DB 781 CGGTGATCATCTGGGCTGAAACAAGATCGTGGATGACAGCCCGTGAAGATCTTG 840  
 QY 841 GACATCCGCGAGAGGCCCAAGAGCCCTTCCGAGATACGTGACCGCTTCTTCAAGACC 900  
 DB 841 GACATCCGCGAGAGGCCCAAGAGCCCTTCCGAGATACGTGACCGCTTCTTCAAGACC 900  
 QY 901 CTGCGGCGGAGAGGCCCAAGAGCCCTTCCGAGATACGTGACCGCTTCTTCAAGACC 960  
 DB 901 CTGCGGCGGAGAGGCCCAAGAGCCCTTCCGAGATACGTGACCGCTTCTTCAAGACC 960  
 QY 961 CAGAAAGCCCAACCCCGAGCTGCAAGACATCTGCGGCTTCTGCGCCCGGCGCACCTTG 1020  
 DB 961 CAGAAAGCCCAACCCCGAGCTGCAAGACATCTGCGGCTTCTGCGCCCGGCGCACCTTG 1020  
 QY 1021 GAGGAGATATGACCGCTGCGAGGCGTGGGCGGCGCGGCGCAAGGCCCGCGTGTG 1080  
 DB 1021 GAGGAGATATGACCGCTGCGAGGCGTGGGCGGCGCGGCGCAAGGCCCGCGTGTG 1080  
 QY 1081 GCGGAGGCGATGAGCCGCGCAAGCGTGAACATCATATGACAGAGCAACTTCAAG 1140  
 DB 1081 GCGGAGGCGATGAGCCGCGCAAGCGTGAACATCATATGACAGAGCAACTTCAAG 1140  
 QY 1141 GCGCCCGGCGCAAGCTCAAGTGTCTTCAACTGCGGAGAGAGGCCCACTCCGCAAGAAC 1200  
 DB 1141 GCGCCCGGCGCAAGCTCAAGTGTCTTCAACTGCGGAGAGAGGCCCACTCCGCAAGAAC 1200

QY 1201 TGCGCGCCCGCCCGCAAGAGGGCTGTGAAGTGGCGCAAGAGGCCACCAAGTGAAG 1260  
 DB 1201 TGCGCGCCCGCCCGCAAGAGGGCTGTGAAGTGGCGCAAGAGGCCACCAAGTGAAG 1260  
 QY 1261 GACTGACCGAGCGCGCAAGGCCCAACTTCTGGCAAGATCTGGCCAGCAGAGGCCCGC 1320  
 DB 1261 GACTGACCGAGCGCGCAAGGCCCAACTTCTGGCAAGATCTGGCCAGCAGAGGCCCGC 1320  
 QY 1321 CCGCGCAACTTCTGCAAGAACCGCAGCGAGCCCGCGCCCACTGTGCCACCGCCCGC 1380  
 DB 1321 CCGCGCAACTTCTGCAAGAACCGCAGCGAGCCCGCGCCCACTGTGCCACCGCCCGC 1380  
 QY 1381 CCGCGAGAGCTTCTGCAAGAACCGCAGCGAGCCCGCGCCCACTGTGCCACCGCCCGC 1440  
 DB 1381 CCGCGAGAGCTTCTGCAAGAACCGCAGCGAGCCCGCGCCCACTGTGCCACCGCCCGC 1440  
 QY 1441 CCGGAGCCCTTACCGGAGCCCTTGAACCGCCCTGCGCAGCCTGTTGCGAGCGGCCCTG 1500  
 DB 1441 CCGGAGCCCTTACCGGAGCCCTTGAACCGCCCTGCGCAGCCTGTTGCGAGCGGCCCTG 1500  
 QY 1501 AGCCAGTAA 1509  
 DB 1501 AGCCAGTAA 1509

RESULT 3  
 US-09-967-464-68  
 / Sequence 68, Application US/09967464  
 / Publication No. US20030138453A1  
 / GENERAL INFORMATION:  
 / APPLICANT: O'Hagan, Derek  
 / APPLICANT: Ottem, Gillis  
 / APPLICANT: Donnelly, John J.  
 / APPLICANT: Polo, John M.  
 / APPLICANT: Barnett, Susan  
 / APPLICANT: Singh, Mamohan  
 / APPLICANT: Umer, Jeffrey  
 / APPLICANT: Dubensky, Jr., Thomas W.  
 / TITLE OF INVENTION: MICROARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS  
 / FILE REFERENCE: PPI6269.004  
 / CURRENT APPLICATION NUMBER: US/09/967,464  
 / PRIOR FILING DATE: 2002-04-11  
 / PRIOR APPLICATION NUMBER: 60/236,105  
 / PRIOR FILING DATE: 2000-09-28  
 / PRIOR APPLICATION NUMBER: 60/315,905  
 / NUMBER OF SEQ ID NOS: 68  
 / SOFTWARE: PatentIn version 3.1  
 / SEQ ID NO: 68  
 / LENGTH: 1509  
 / TYPE: DNA  
 / ORGANISM: Human immunodeficiency virus type 1  
 US-09-967-464-68

Query Match 99.7%; Score 1504.2; DB 10; Length 1509;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1497; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGGCGCAGCATCTGCGCGGCGAGAAAGTGAACAAGTGGAGAAATCCGC 60  
 DB 1 ATGGGCGCCCGGCGCAGCATCTGCGCGGCGAGAAAGTGAACAAGTGGAGAAATCCGC 60  
 QY 61 CTGCGCCCGGCGGCAAGACATACTGCTGAAGCACTGTGTGGGCCAGCGCGAG 120  
 DB 61 CTGCGCCCGGCGGCAAGACATACTGCTGAAGCACTGTGTGGGCCAGCGCGAG 120  
 QY 121 CTGAGAGGCTTGGCTTAACCCCGGCTGCTGAGAGCCGCGGAGGCTGCAAGCAATC 180  
 DB 121 CTGAGAGGCTTGGCTTAACCCCGGCTGCTGAGAGCCGCGGAGGCTGCAAGCAATC 180  
 QY 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGCGCAAGAGAGCTGCGCAGCTGTACAAC 240  
 DB 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGCGCAAGAGAGCTGCGCAGCTGTACAAC 240



```

Db      361 GCGGAGCCCTAAGCCGAGACCCCTGACCGCCTTGCGAGACTGTTCGGAGCGGCCCTTG 420
Qy      421 CACGAGGCGATAGCCCCCGCACCTGAAAGCTTGAGTGAAGTATCGAGAGAAAGCC 480
Db      421 CACGAGGCGATAGCCCCCGCACCTGAAAGCTTGAGTGAAGTATCGAGAGAAAGCC 480
Qy      481 TTGAGCGCCGAGTGTATCCCAATGTTCAACGCGCTGAGAGAGAGGCGCCACCCCGAGAG 540
Db      481 TTGAGCGCCGAGTGTATCCCAATGTTCAACGCGCTGAGAGAGAGGCGCCACCCCGAGAG 540
Qy      541 CTGAAACAGCATGTTGAACAACGCTGGCGCGCACAGCGCGCATGAGATGCTGAAGAC 600
Db      541 CTGAAACAGCATGTTGAACAACGCTGGCGCGCACAGCGCGCATGAGATGCTGAAGAC 600
Qy      601 ACCATCAACGAGAGAGCGCCGCAAGTGGAGACCGCTGACCCCGTGAAGCGCGCCCTG 660
Db      601 ACCATCAACGAGAGAGCGCCGCAAGTGGAGACCGCTGACCCCGTGAAGCGCGCCCTG 660
Qy      661 GCGCGCGCGAGATGCGGACCCCGCGGACGACATGCGCGCGCGCCACGACACCTG 720
Db      661 GCGCGCGCGAGATGCGGACCCCGCGGACGACATGCGCGCGCGCCACGACACCTG 720
Qy      721 CAGAGACAGATGCGCTGATGACACAGCAACCCCGCGTGCCTGGGCGAGCATCTAACG 780
Db      721 CAGAGACAGATGCGCTGATGACACAGCAACCCCGCGTGCCTGGGCGAGCATCTAACG 780
Qy      781 CGGTGATCATCTGGGCTGAAACAAGATGTCGAGATGTACAGCCCGTGAACATCTTG 840
Db      781 CGGTGATCATCTGGGCTGAAACAAGATGTCGAGATGTACAGCCCGTGAACATCTTG 840
Qy      841 GACATCCGCGAGAGCGCCGAAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTTCAAGAC 900
Db      841 GACATCCGCGAGAGCGCCGAAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTTCAAGAC 900
Qy      901 CTGCGCGCGAGACAGGCGCACCCAGGACGTGAAGAACTGATATCGAGACCCCTGCTGTG 960
Db      901 CTGCGCGCGAGACAGGCGCACCCAGGACGTGAAGAACTGATATCGAGACCCCTGCTGTG 960
Qy      961 CAGAAACGCAACCCCGCATGCAAGACATCTGCGCGCTTCTGCGCGCGCGCGCACCTG 1020
Db      961 CAGAAACGCAACCCCGCATGCAAGACATCTGCGCGCTTCTGCGCGCGCGCGCACCTG 1020
Qy      1021 GAGGAGATGATGACCGCTGCGAGGCGCTGGCGCGCGCGCGCGCACAGGCGCGCGCTG 1080
Db      1021 GAGGAGATGATGACCGCTGCGAGGCGCTGGCGCGCGCGCGCGCACAGGCGCGCGCTG 1080
Qy      1081 GCGGAGGCGATGAGCGAGCGCAACAGCTGAACATCATGATGACAGAGCAACTTCAAG 1140
Db      1081 GCGGAGGCGATGAGCGAGCGCAACAGCTGAACATCATGATGACAGAGCAACTTCAAG 1140
Qy      1141 GCGCGCGCGCGCAACGTCAGAGTCTTCAACTGCGGCAAGAGGCGCACTGCCAAGAAC 1200
Db      1141 GCGCGCGCGCGCAACGTCAGAGTCTTCAACTGCGGCAAGAGGCGCACTGCCAAGAAC 1200
Qy      1201 TGGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGGCGCAAGAGGCGCGCAAGTGAAG 1260
Db      1201 TGGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGGCGCAAGAGGCGCGCAAGTGAAG 1260
Qy      1261 GACTGACCGAGCGCGAGCGCAACTTCTGAGGCAAGATCTGGCGCAAGAGGCGCGC 1320
Db      1261 GACTGACCGAGCGCGAGCGCAACTTCTGAGGCAAGATCTGGCGCAAGAGGCGCGC 1320
Qy      1321 CCGCGCAACTTCTGAGAAACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db      1321 CCGCGCAACTTCTGAGAAACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy      1381 CCGCGCAAGAGCTTCCGCTTGAAGAGAACACCCCGCGCGCGCGCGCGCGCGCGCG 1440
Db      1381 CCGCGCAAGAGCTTCCGCTTGAAGAGAACACCCCGCGCGCGCGCGCGCGCGCGCG 1440
Qy      1441 CCGGAGCCCTAAGCGAGCGCGCTGACCGCGCTGCGAGCGCTTTCGCGAGCGCGCGCG 1500
Db      1441 CCGGAGCCCTAAGCGAGCGCGCTGACCGCGCTGCGAGCGCTTTCGCGAGCGCGCGCG 1500

```

```

Db      1441 GCGGAGCCCTAAGCCGAGACCCCTGACCGCCTTGCGAGACTGTTCGGAGCGGCCCTTG 1500
Qy      1501 AGCCAGTAA 1509
Db      1501 AGCCAGTAA 1509

RESULT 5
US-09-899-575-51
; Sequence 51, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: Van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Type C
; US-09-899-575-51

Query Match      86.0%; Score 1297.6; DB 10; Length 1494;
Best Local Similarity 92.7%; Pred. No. 1,4e-271;
Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;

Qy      1 ATGGGCGCGCGCGCGCGCGATCTTGCGCGCGCGCGAGAAAGTGAACAAAGTGGAGAGATCCGC 60
Db      1 ATGGGCGCGCGCGCGCGCGCGATCTTGAGCGCGCGCGAGAGTGAACAAAGTGGAGAGATCCGC 60
Qy      61 CTGCGCGCGCGCGCGAGAAACATACATGCTGAAGACCTGTGTGGGCGACCGCGGAG 120
Db      61 CTGCGCGCGCGCGCGAGAAACATACATGCTGAAGACCTGTGTGGGCGACCGCGGAG 120
Qy      121 CTGAGAGGCTTTCGCTGTAACCCCGCGCTGCTGAGAACCGCGAGGCTGCAAGCAGATC 180
Db      121 CTGAGAGGCTTTCGCTGTAACCCCGCGCTGCTGAGAACCGCGAGGCTGCAAGCAGATC 180
Qy      181 ATGAAGACCTGACAGCGCGCGCTGAGAACCGGCAACCGAGAGCTGCGAGCCTGTACAC 240
Db      181 ATGAAGACCTGACAGCGCGCGCTGAGAACCGGCAACCGAGAGCTGCGAGCCTGTACAC 240
Qy      241 ACCGTGGCGACCTGTATGCTGTGACCGCGCGCATCGAGTCCCGAGAACCAAGAGGCC 300
Db      241 ACCGTGGCGACCTGTATGCTGTGACCGCGCGCATCGAGTCCCGAGAACCAAGAGGCC 300
Qy      301 CTGGAACAAGTGAAGAGAGAGAGCAACAAGTCCAGCGAGAGAGACCCAGAGGAGAG 360
Db      301 CTGGAACAAGTGAAGAGAGAGAGCAACAAGTCCAGCGAGAGAGAGACCCAGAGGAGAG 360
Qy      361 GCGGAGCGCAAGTGAAGCGCAAGTCCATCTGTGCAAACTTGACAGGCGCGAGTGTG 420
Db      361 GCGGAGCGCAAGTGAAGCGCAAGTCCATCTGTGCAAACTTGACAGGCGCGAGTGTG 420
Qy      421 CACGAGGCGATCAAGCGCGCGCGCAACCTGAAGCGCTGGGAGAGTGAAGAGAGAGCC 480
Db      421 CACGAGGCGATCAAGCGCGCGCGCAACCTGAAGCGCTGGGAGAGTGAAGAGAGAGCC 480
Qy      481 TTGAGCGCCGAGTGTATCCCAATGTTCAACGCGCTGAGAGAGAGGCGCCACCCCGAGAG 540
Db      481 TTGAGCGCCGAGTGTATCCCAATGTTCAACGCGCTGAGAGAGAGGCGCCACCCCGAGAG 540

```

```

QY 541 CTGAACACGATGTTGAACAACCGTGGGCGGCAACCAAGCCGCAATGCAGATGCTGAAGAC 600
DB 541 CTGAACACCATCTGTGAACAACCGTGGGCGGCAACCAAGCCGCAATGCAGATGCTGAAGAC 600
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTTGACCCGCTGCAAGGCCGCGCCGCTG 660
DB 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCCGCAACCCGCTGCAAGGCCGCGCCGCTG 660
QY 661 GCGCCGCGCGAGATGGCGCAACCCCGCGCGAGAGATGCGCGGCAACCAAGCCGCTG 720
DB 661 GCGCCGCGCGAGATGGCGCAACCCCGCGCGAGAGATGCGCGGCAACCAAGCCGCTG 720
QY 721 CAGAGACGATCGCTGATGACACAGACAACCCCGCTGCGCGCGAGATCTTCAAG 780
DB 721 CAGAGACGATCGCTGATGACACAGACAACCCCGCTGCGCGAGATCTTCAAG 780
QY 781 CGGTGATCATCTGTGGCTGAAACAAGATCTGTGGATGACAGCCCGTGAACATCTG 840
DB 781 CGGTGATCATCTGTGGCTGAAACAAGATCTGTGGATGACAGCCCGTGAACATCTG 840
QY 841 GACATCCGCGAGGCGCCCAAGAGCCCTTCCGGAATACGTGAGACCGCTTTCAGAGAC 900
DB 841 GACATCAAGAGGCGCCCAAGAGCCCTTCCGGAATACGTGAGACCGCTTTCAGAGAC 900
QY 901 CTGCGCGCGAGCGCAAGCCCAAGAGCTGAAGAATGATGACCGAGACCGCTGCTG 960
DB 901 CTGCGCGCGAGCGCAAGCCCAAGAGCTGAAGAATGATGACCGAGACCGCTGCTG 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGTGGCTTCTGCGCCCGCGCGCAACCTG 1020
DB 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGTGGCTTCTGCGCCCGCGCGCAACCTG 1020
QY 1021 GAGGAGATGATACCGCTGCAAGGCGGTGGGCGCGCGCGCAAGAGCCCGCGCTG 1080
DB 1021 GAGGAGATGATACCGCTGCAAGGCGGTGGGCGCGCGCGCAAGAGCCCGCGCTG 1080
QY 1081 GCCGAGCGATGAGCCAGGCGCAACAGCTGGAACATCATGATGACGAAGAGCAATTCAG 1140
DB 1081 GCCGAGCGCATGAGCCAGGCGCAACAGCTGGAACATCATGATGACGAAGAGCAATTCAG 1140
QY 1141 GCGCCCGCGCGCAACCTGCAAGTGTCTTCACTGTGGCGCAAGAGGCGCAATTCGCGCAAG 1200
DB 1141 GCGCGCAACCGGATCATCAAGTGTCTTCACTGTGGCGCAAGAGGCGCAATTCGCGCAAG 1200
QY 1201 TGCCTGCGCGCGCGCAAGAGGCGGTGGAGTGGCGCAAGAGGCGCAACGATGAG 1260
DB 1201 TGCCTGCGCGCGCGCGCAAGAGGCGGTGGAGTGGCGCAAGAGGCGCAACGATGAG 1260
QY 1261 GACTGACGAGAGCGCGCAACCTTCTGTGGCGCAAGATCTGTGGCGCAAGAGGCGCG 1320
DB 1261 GACTGACGAGAGCGCGCAACCTTCTGTGGCGCAAGATCTGTGGCGCAAGAGGCGCG 1320
QY 1321 CCGCGCACTTCTGTGGAACCGCGAGAGC---CGCGCGCGCGCAACCGTGCACCGCG 1377
DB 1321 CCGCGCACTTCTGTGGAACCGCGAGAGC---CGCGCGCGCGCAACCGTGCACCGCG 1377
QY 1378 CCGCGCGCGAGAGCTTCCGCTTCAAGAGAGACCAACCCCGTGTGCGCGCAAGAGAG 1437
DB 1378 CCGCGCGCGAGAGCTTCCGCTTCAAGAGAGACCAACCCCGTGTGCGCGCAAGAGAG 1437
QY 1438 GACCGGAGACCTTACCGCGAGCGCGCTGACCGCGCTGTTCCGCGAGCGCGCG 1497
DB 1438 -----CGGAGCGCGCTGACCGCGCTGTTCCGCGAGCGCGCG 1482
QY 1498 CTGAGCCAGTAA 1509
DB 1483 CTGAGCCAGTAA 1494

```

RESULT 6  
US-09-899-575-99  
; Sequence 99, Application US/09899575

```

; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnet, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gag_TIV2_C_Zhopt
US-09-899-575-99

Query Match      85.8%; Score 1294.8; DB 10; Length 1491;
Best Local Similarity 92.7%; Pred. No. 5.8e-271;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;

QY 1 ATGGGCGCGCGCGCGCAACATCTGTGGCGCGCGAGAAAGCTGGAACATGCGGAATCCGC 60
DB 1 ATGGGCGCGCGCGCGCAACATCTGTGGCGCGCGCAAGCTGGAACATGCGGAATCCGC 60
QY 61 CTGCGCGCGCGCGCGCAAGACATCACTGTAAGACCTGATGTGGCGCGCGCGAG 120
DB 61 CTGCGCGCGCGCGCGCGCAAGACATCACTGTAAGACCTGATGTGGCGCGCGCGAG 120
QY 121 CTGAGGCGCTTGGCGCTTGAACCCCGCGCTGTGAGACCGCGAGGCTGCAAGCATC 180
DB 121 CTGAGGCGCTTGGCGCTTGAACCCCGCGCTGTGAGACCGCGAGGCTGCAAGCATC 180
QY 181 ATGAAGCAGCTGCAAGCCCGCGCTGAGACCGCGACCGGAGAGCTGGGAGCTGTAAAC 240
DB 181 ATGAAGCAGCTGCAAGCCCGCGCTGAGACCGCGACCGGAGAGCTGGGAGCTGTAAAC 240
QY 241 ACCGTGGCGCAACCTGTATCTGCGTGTGACCGCGCATCGAGTCCGCGACCAAGAGGCG 300
DB 241 ACCGTGGCGCAACCTGTATCTGCGTGTGACCAAGGATCGAGTCCGCGACCAAGAGGCG 300
QY 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTCCAGCAAGAAACCCAGAGCGCAAGAG 360
DB 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTCCAGCAAGAAACCCAGAGCGCAAGAG 360
QY 361 GCGGAGCGCAAGTGAAGCAACAATCACTGTCGAACCTGTGAGGCGCGCATATGTTG 420
DB 361 GCGGAGCGCAAGTGAAGCAACAATCACTGTCGAACCTGTGAGGCGCGCATATGTTG 420
QY 421 CACGAGCGCATAGCGCGCGCAACCTGTAACGCTGTGAGAGTATGAGAGAGAGCGC 480
DB 421 CACGAGCGCATAGCGCGCGCAACCTGTAACGCTGTGAGAGTATGAGAGAGAGCGC 480
QY 481 TTGAGCGCGAGTATCCCATTTCAACCGCTGTGAGAGGCGCGCGCGCGCGCGAGAG 540
DB 481 TTGAGCGCGAGTATCCCATTTCAACCGCTGTGAGAGGCGCGCGCGCGCGCGAGAG 540
QY 541 CTGAACAAGATGTTGAACAACCTGTGGCGCGCGCAACAGCGCGCGCATGCAATCTGAAGAC 600
DB 541 CTGAACAAGATGTTGAACAACCTGTGGCGCGCGCAACAGCGCGCGCATGCAATCTGAAGAC 600
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTGACCGCGCTGACAGCGCGCGCGCTG 660
DB 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTGACCGCGCTGACAGCGCGCGCGCTG 660
QY 661 GCGCCGCGCGAGATGGCGCAACCCCGCGCGAGAGATGCGCGGCAACCAAGCCGCTG 720
DB 661 GCGCCGCGCGAGATGGCGCAACCCCGCGCGAGAGATGCGCGGCAACCAAGCCGCTG 720

```





```

Db      781 AAGCGGTGATCATCTCTGGGCTCTGAACAAGATCGTGGCATGTAGACCCCGTGAGCATC 840
QY      838 CTGGAATCCGCGCAGAGGCCCCCAAGAGGAGCCCTTCCGGAGCTAGTGGACCGCTTCTTCAAG 897
Db      841 CTGGAATCAAGACAGAGGCCCCCAAGAGGAGCCCTTCCGGAGCTAGTGGACCGCTTCTTCAAG 900
QY      898 ACCCTGCGCGCGAGAGGAGCCCAAGAGGAGCTGGAAGAACTGTATGACCGAGACCCCTGTCTG 957
Db      901 ACCCTGCGCGCGAGAGGAGCCCAAGAGGAGCTGGAAGAACTGTATGACCGAGACCCCTGTCTG 960
QY      958 GTGCAAGAACCCCAACCCCGAATCTGCGAGACCATCTGCGGCTCTGCGGCCCGCGCCAGC 1017
Db      961 GTGCAAGAACCCCAACCCCGAATCTGCGAGACCATCTGCGGCTCTGCGGCCCGCGCCAGC 1020
QY      1018 CTGGAAGAGATGATGACCGCTGCGAGGAGGCTGAGGAGGCGCCCGGACCAAGGCGCGCGT 1077
Db      1021 CTGGAAGAGATGATGACCGCTGCGAGGAGGCTGAGGAGGCGCCCGGACCAAGGCGCGCGT 1080
QY      1078 CTGCGCGAGGCGATGAGCCAGGCGCAACAGCTGAACATCATGATGAGCAAGAGCAATTC 1137
Db      1081 CTGCGCGAGGCGATGAGCCAGGCGCAACAC---AGGTGATGATGAGCAAGAGCAATTC 1137
QY      1138 AAGGCGCGCGCGCAAGCTCAAGTCTTCAACTGCGGCGCAAGAGGCGCAATCTGCGCAAG 1197
Db      1138 AAGGCGCGCGCGCGATCTCAAGTCTTCAACTGCGGCGCAAGAGGCGCAATCTGCGCGC 1197
QY      1198 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGTGCGCAAGAGGCGCGCAAGAGT 1257
Db      1198 AACTGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGTGCGCAAGAGGCGCGCAAGAGT 1257
QY      1258 AAGGACTGACCGAGAGCGCGCAACCTTCTGAGGCAAGATCTGCGCGCAAGCGCAAGGCG 1317
Db      1258 AAGGACTGACCGAGAGCGCGCAACCTTCTGAGGCAAGATCTGCGCGCAAGCGCAAGGCG 1317
QY      1318 CGCGCGCGCGCAATCTCTGAGAAACCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db      1318 CGCGCGCGCGCAATCTCTGAGAAACCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCG 1359
QY      1378 CGCGCGCGCGAGAGCTTCTGAGAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db      1380 CGCGCGCGCGAGAGCTTCTGAGAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1419
QY      1438 GACCGCGAGCCCTTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db      1420 G-----ACCGCGAGAGCCCTGACCGAGCTGGAAGAGCTGTGCGGCAAGCGCGCG 1467
QY      1498 CTGAGCCCACTAA 1509
Db      1468 CTGAGCCCACTAA 1479

```

## RESULT 8

```

US-09-899-575-3
; Sequence 3, Application US/0989575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnette, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR APPLICATION NUMBER: 09/475,704
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110965
US-09-899-575-3

```

```

Query Match      84.6%; Score 1276.8; DB 10; Length 1479;
Best Local Similarity 92.5%; Pred. No. 4,5e-267;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

```

```

QY      1 ATGGGCGCGCGCGCGAGATCTGCGGCGGAGAAAGTGGACAAAGTGGAGAAATCCGCG 60
Db      1 ATGGGCGCGCGCGCGAGATCTGCGGCGGCGGAGAAAGTGGAGCGCTGGAGCGCATCCGCG 60
QY      61 CTGCGCGCGCGCGGCAAGAACATCACTAGTCTGAAGCACTGTGTGTGGGCGAGCGCGAG 120
Db      61 CTGCGCGCGCGCGGCAAGAACATGCTACTATGATGAAGCACTGTGTGTGGGCGAGCGCGAG 120
QY      121 CTGGAAGGCTTGTGCTGAAACCCCGGCTGTGAGAGACCGCGGAGGCTGCAAGCATATC 180
Db      121 CTGGAAGAGTTGTGCTGAAACCCCGGCTGTGAGAGACCGAGGCTGCAAGCATATC 180
QY      181 ATGAAGGAGCTGCAAGCGCGCGCTGCAAGACCGGAGAGCTGCGGAGCTGTGACATAC 240
Db      181 ATCGCGCGAGCTGCAAGCGCGCGCTGCAAGACCGGAGAGCTGGAAGAGCTGTGACATAC 240
QY      241 ACCGTGCGCACCCCTGTACTGTGCTGCAAGCGCGCATCGAGTCCGCGACACCAAGAGGCGC 300
Db      241 ACCGTGCGCACCCCTGTACTGTGCTGCAAGAGAAATCGAGTCCGCGACACCAAGAGGCGC 300
QY      301 CTGGAAGAGATGAGAGAGAGCAAGCAAGTCCAGCAAGAAAGACCCAGAGCGCAAGAGAG 360
Db      301 CTGGAAGAGATGAGAGAGAGCAAGCAAGTCCAGCAAGAAAGTCCAGCAAGAGCGCAAGAG 360
QY      361 GCGGAG---GGCAAGTGTAGCAAGAACCTATCTGTGCAAGTCTGCAAGAGCGCGCGAGT 417
Db      361 GCGGAGAGGCGAGAGTGTAGCAAGAACCTATCTGTGCAAGTCTGCAAGAGCGCGCGAGT 420
QY      418 GTGCAACAGGCGCATAGCGCGCGGACCCCTGGAACCGCTGTGAGTGTGAGAGAGAG 477
Db      421 GTGCAACAGGCGCATAGCGCGCGGACCCCTGGAACCGCTGTGAGTGTGAGAGAGAG 480
QY      478 GCTTTACGCGCGAGAGTATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCAACCCCGAG 537
Db      481 GCTTTACGCGCGAGAGTATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCAACCCCGAG 540
QY      538 GACCTGAACAGAGTGTGAACACCGTGGGCGGCGCAAGCGCGCGCATATGATGTAGAG 597
Db      541 GACCTGAACAGAGTGTGAACACCGTGGGCGGCGCAAGCGCGCGCATATGATGTAGAG 600
QY      598 GACACCATCAACAGAGAGCGCGCGAGTGGAGACCGCTGCAACCCCGGTGCAAGCGCGCGCG 657
Db      601 GACACCATCAACAGAGAGCGCGCGAGTGGAGACCGCGGTGCAACCCCGGTGCAAGCGCGCGCG 660
QY      658 GTGCGCGCGCGCGAGATCGCGAGACCCCGCGGCGAGCGCATCGCGCGCGCAAGCAAGC 717
Db      661 ATCGCGCGCGCGAGATCGCGAGACCCCGCGGCGAGCATCGCGCGCGCAAGCAAGC 720
QY      718 CTGGAAGAGAGATGTGCTGTGATGACCAAGCAACCCCGGTGCGCGGTGGAGATCTAC 777
Db      721 CTGGAAGAGAGATGTGCTGTGATGACCAAGCAACCCCGGTGCGCGGTGGAGATCTAC 780
QY      778 AAGCGGTGATCATCTGTGGCGCTGAACAAAGATGTGCGGATGTACAGCGCGGTGAGCATC 837
Db      781 AAGCGGTGATCATCTGTGGCGCTGAACAAAGATGTGCGGATGTACAGCGCGGTGAGCATC 840
QY      838 CTGGAATCCGCGCAGAGGCGCGCAAGAGGCTTCCGCGCATGAGAGCGGCTTCTTCAAG 897
Db      841 CTGGAATCAAGAGGCGCGCAAGAGGCTTCCGCGCATGAGAGCGGCTTCTTCAAG 900
QY      898 ACCCTGCGCGCGAGAGCGCAACCAAGAGAGCTGAAGAACTGTGATGACCGAGACCTGTGTG 957
Db      901 ACCCTGCGCGCGAGAGAGCAACCAAGAGAGTGAAGAACTGTGATGACCGAGACCTGTGTG 960

```



[illegible]

	Query Match	Similarity	Score	1275.2;	DB 15;	Length	3162;
	Best Local	Similarity	92.5%;	Pred.	No. 9e-26;		
	Matches	1998;	Conservative	0;	Mismatches	76;	Indels
							Gaps
							36;
							Gaps
							4;
Qy	1	ATGGGCGCCGCGCAGCATCTCTGCGCGCGCGAAGACTGGA	CAAGTGGAGAAATCTCGC	60			
Db	7	ATGGGCGCCGCGCAGCATCTCTGCGCGCGCGAAGCTGAGACGCTTGGAGACGATCCGC	66				

QY	61	CTGGCCCCGGGCGAAGCACTACATGTCGTAAGCACTGCTGTGGGCCAGCCGCGAG	120
Db	67	CTGGCCCCGGGCGAAGAGTGCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG	126
QY	121	CTGGAGGGCTTTCGCTCTGAACCCCGGGCTGTGAGACCGCGCGAGGGCTGCAAGCATTC	180
Db	127	CTGGAGGAAGTTTCGCTCTGAACCCCGGGCTGTGAGAGCAAGAGGGCTGCAAGCATTC	186
QY	181	ATGAAGCAGCTGCAGCCCCCTTGACACCGGCACCGAGAGACTGCGACTGTTCATAC	240
Db	187	ATCGCCCACTGCAGCCCCCGCTTGACACCGGCAGCGAGGAGCTGAAGAGCTGTTCATAC	246
QY	241	ACCGTGGCCACCTGTCTACTGCGTGACGCGCGCATTCAGAGTCCCGGCAACCAAGGAGGCC	300
Db	247	ACCGTGGCCACCTGTCTACTGCGTGACGAGAAAGATCGAGTCCCGGCAACCAAGGAGGCC	306
QY	301	CTGGACAAGATCGAGAGAGCAAGACAAAGTCCGACGAAAGACCAGCAGGCCCAAGAG	360
Db	307	CTGGACAAGATCGAGAGAGCAAGACAAAGTCCGACGAGAGATTCAGCAGGCCCAAGAGCC	366
QY	361	GCCGAC---GGCAAGGTGAGGCAAGAACTAACCCATTCGTGAGAACTCTGCAAGGCCAAGT	417
Db	367	GCCGACAAAGGCAAGGTGAGGCAAGAACTAACCCATTCGTGAGAACTCTGCAAGGCCAAGT	426
QY	418	GTGCACACGAGCATCAGCGCCCCGCAACCTTGAAACCGCTGGTGAAGGTATCGAGAGAG	477
Db	427	GTGCACACGAGCATCAGCGCCCCGCAACCGCTGGTGAAGGTATCGAGAGAGAG	486
QY	478	GCTTTAGCCCCCGAGGTGATCCCATGTTTCAACCGCTCTGAGGAGAGGCGCCACCCCCAG	537
Db	487	GCTTTAGCCCCCGAGGTGATCCCATGTTTCAACCGCTCTGAGCGAGGGCGCCACCCCCAG	546
QY	538	GACCTGAACAGATGTTGAACAACGGTGGGGCGGCAACAGGGCGGCATGCAGATGCTGAAG	597
Db	547	GACCTGAACAGATGTTGAACAACGGTGGGGCGGCAACAGGGCGGCATGCAGATGCTGAAG	606
QY	598	GACACCATCAACGAGAGAGCGCGCGAGTGGAGCCGCTTGCAACCCCGTCAAGGCCGAGCCC	657
Db	607	GACACCATCAACGAGAGAGCGCGCGAGTGGAGCCGCGTGCACCCCGTCAAGGCCGAGCCC	666
QY	658	GTGGCCCCCGGCGCAGATTCGCGGACCCCGCGCAGCGCATCTGCGCGGCGCCACAGCAC	717
Db	667	ATCGCCCCCGGCGCAGATTCGCGGAGCCCGCGCAGCGCATCTGCGCGGCGCCACAGCAC	726
QY	718	CTGCAGAGAGCATGCTGCTGATGACACGACAACCCCCCGTGGCCGTGGGCGACATCTAC	777
Db	727	CTGCAGAGAGCATGCTGCTGATGACACGACAACCCCCCATCTCCGTGGGCGACATCTAC	786
QY	778	AAGCGGTGATCATCTTGGGCTTGAAACAAGATCGTCCGATGTACAGCCCCGTGAGCATC	837
Db	787	AAGCGGTGATCATCTTGGGCTTGAAACAAGATCGTCCGATGTACAGCCCCGTGAGCATC	846
QY	838	CTGCACATCCCGCAGGGGCGCCCAAGAGGCCCTTCGCGCATCTACGTGAGACCGCTTCTTCAAG	897
Db	847	CTGCACATCAACAGGGGCGCCCAAGAGGCCCTTCGCGCATCTACGTGAGACCGCTTCTTCAAG	906
QY	898	ACCCTGCGCGCGAGCAGGCCACCCAGGACGTGAAGAACTGATGACCGAGACCTCTGCTG	957
Db	907	ACCCTGCGCGCGCGAGCAGGCCACCCAGGAGGTGAAGAACTGATGACCGAGACCTCTGCTG	966
QY	958	GTGCAGAACGCGAACCCCTGACTGCAAGCATCTTGGCGCTCTTGGGCCCGGCGCCAC	1017
Db	967	GTGCAGAACGCGAACCCCTGACTGCAAGCATCTTGGCGCTCTTGGGCCCGGCGCCAC	1026
QY	1018	CTGGAGGAGATGATGACGCGCTGCGCAAGGCGGTGGGCGGCCCGGCAACAAGCCGCGCTG	1077
Db	1027	CTGGAGGAGATGATGACGCGCTGCGCAAGGCGGTGGGCGGCCCGGCAACAAGCCGCGCTG	1086
QY	1078	CTGGCCGAGGCGATGAGCCAGGCCAACAAGCGTGAACATCATGTGCAAGAGCACTTC	1137
Db	1087	CTGGCCGAGGCGATGAGCCAGGCCAACAAC---AGCGTGAAGATGCAAGAGCAACTTC	1147



```
QY 1318 CCCCCCGCAACTTCCTGCGAAGACCGAGCGAGCCCGCGCCGCCACCGCTGCGCCACCGCC 1377
    |||||
Db 1324 CCCCCCGCAACTTCCTGCGAAGACCG-----CCCGAGGCCACCGCC 1365
QY 1378 CCCCCCGCGAAGACTTCCTGCTTCGAGAGACCAACCCCGCCGCCAAGAGAGAGCCCAAG 1437
    |||||
Db 1366 CCCCCCGCGAAGACTTCCTGCTTCGAGAGACCAACCCCGCGCGCAAGAGAGAGAGCAAG 1425
QY 1438 GACCCGAGACCCCTTACCGCGAGCCCTTGACCGCCCTCGCGACCGCTTTCGCGAGCGGCC 1497
    |||||
Db 1426 G-----ACCGGAGACCTTGACCGAGCTTGAAAGCCCTGTTCCGCAACGACCC 1473
QY 1498 CTGAGCCAGTAA 1509
    |||||
Db 1474 CTGAGCCAGAAA 1485

RESULT 11
US-10-190-435-19
; Sequence 19, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmutTatRevNef_C
US-10-190-435-19

Query Match      84.5%; Score 1275.2; DB 15; Length 4419;
Best Local Similarity 92.5%; Pred. No. 9,6e-267;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
```

```
QY 418 GTGCAACAGAGCCATCAAGCCCGCGACCTTGAAACGCTTGAGTGAAGTATCGAGAGAAAG 477
    |||||
Db 427 GTGCAACAGAGCCATCAAGCCCGCGACCTTGAAACGCTTGAGTGAAGTATCGAGAGAAAG 486
QY 478 GCTTCAGCCCGAGAGTGTATCCCAATGTTTCAACCGGCTTGAGAGAGAGGCGCCACCCCGAG 537
    |||||
Db 487 GCTTCAGCCCGAGAGTGTATCCCAATGTTTCAACCGGCTTGAGAGAGAGGCGCCACCCCGAG 546
QY 538 GACTGAACACAGATGTTGAACAACGCTGGAGCGGCGACACAGGCCCGCATGAGATGCTGAAG 597
    |||||
Db 547 GACTGAACACAGATGTTGAACAACGCTGGAGCGGCGACACAGGCCCGCATGAGATGCTGAAG 606
QY 598 GACACCATCAACAGAGAGCGCGCGAGTGGAGACCGCTTGACACCCGTGAGCGCCGCGCCCG 657
    |||||
Db 607 GACACCATCAACAGAGAGCGCGCGAGTGGAGACCGCTTGACACCCGTGAGCGCCGCGCCCG 666
QY 658 GTGGCCCGCGGCAATGCGGAGACCCCGCGGCAACGACATCGCGGCGCCACCGACGACCC 717
    |||||
Db 667 ATGCCCCCGGCGAATGCGGAGCGAGCCCGCGGCAACGACATCGCGGCGCCACCGACGACCC 726
QY 718 CTGAGAGAGAGATGCGCTGATGACACGAAACCCCGCTGCGCCGCGGCGACATCTAC 777
    |||||
Db 727 CTGAGAGAGAGATGCGCTGATGACACGAAACCCCGCATCCCGTGGCGGACATCTAC 786
QY 778 AAGCGGTGATCATCTGCGCTGAAACAAGATGTCGAGATGTACAGCCCGTGAGCATC 837
    |||||
Db 787 AAGCGGTGATCATCTGCGCTGAAACAAGATGTCGAGATGTACAGCCCGTGAGCATC 846
QY 838 CTGACATTCGCGAGAGGCGCCCAAGAGAGCCCTTCGCGCACTAGTGAACCGCTTCTTCAAG 897
    |||||
Db 847 CTGACATTCGCGAGAGGCGCCCAAGAGAGCCCTTCGCGCACTAGTGAACCGCTTCTTCAAG 906
QY 898 ACCCTGCGCGCGAGACAGAGCCACCGAGACGTGAAGACTGATGATGACGAGACCCCTGCTG 957
    |||||
Db 907 ACCCTGCGCGCGAGACAGAGCCACCGAGAGGTGAAGACTGATGATGACGAGACCCCTGCTG 966
QY 958 GTGCAAAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTTGCGCCCGCGCCACCC 1017
    |||||
Db 967 GTGCAAAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTTGCGCCCGCGCCACCC 1026
QY 1018 CTGAGAGAGATGATACCGGCTGCGAGAGGCGGTGGGCGCCCGGCGCAAGGCCCGCGTGG 1077
    |||||
Db 1027 CTGAGAGAGATGATACCGGCTGCGAGAGGCGGTGGGCGCCCGGCGCAAGAGGCCCGCGTGG 1086
QY 1078 CTGCGCGAGCGATGAGCGAGGCGCAACGCGTGAACATCATGATGACAGAAAGCAACTTC 1137
    |||||
Db 1087 CTGCGCGAGCGATGAGCGAGGCGCAACCGCAACCGCAACCGCAACCGCAACCGCAACTTC 1143
QY 1138 AAGGCGCGCGCGGCGCAACGTCATGCTTCAACTGCGCGCAAGAGAGGCGCAACCGCAAG 1197
    |||||
Db 1144 AAGGCGCGCGCGGCGCATGTCATGCTTCAACTGCGCGCAAGAGAGGCGCAACCGCGCG 1203
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGTGCGGCAAGAGAGGCGCAACGATG 1257
    |||||
Db 1204 AACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGTGCGGCAAGAGAGGCGCAACGATG 1263
QY 1258 AAGGACTGACCGAGCGCGCAAGCGCAACTTCGCGGCAAGATCTGCGCGCAAGCGCAAGAGG 1317
    |||||
Db 1264 AAGGACTGACCGAGCGCGCAAGCGCAACTTCGCGGCAAGATCTGCGCGCAAGCGCAAGAGG 1323
QY 1318 CCGCCCGGCAACTTCTCTGCAAGACCGCAAGCGAGCGCGCGCCCGCATGCTGCCACCGCC 1377
    |||||
Db 1324 CCGCCCGGCAACTTCTCTGCAAGCGG-----CCCGAGGCCACCGCC 1365
QY 1378 CCCCCCGCGAAGACTTCCTGCTTCGAGAGACCAACCCCGCCGCCAAGAGAGAGCCCAAG 1437
    |||||
Db 1366 CCCCCCGCGAAGACTTCCTGCTTCGAGAGACCAACCCCGCGCGCAAGAGAGAGAGCAAG 1425
QY 1438 GACCGGAGCCCTTACCGGAGCGCCCTTGACCGCCCTTGACCGCTTTCGCGAGCGGCCCG 1497
    |||||
Db 1426 G-----ACCGGAGACCTTGACCGAGCTTGAAAGCCCTGTTCCGCAACGACCC 1473
```



```

; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 4483
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-10-190-435-35

```

Query Match 84.5%; Score 1275.2; DB 15; Length 4483;

Best Local Similarity 92.5%; Pred. No. 9.6e-267;

Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

QY 1 ATGGGCGCCCGCCGACATCTCGCGCGGAGAGCTGGACAAAGTGGAGAGATCCCG 60
DB 3005 ATGGGCGCCCGCCGACATCTCGCGCGGAGAGCTGGACAAAGTGGAGAGATCCCG 3064
QY 61 CTGGCGCCCGCCGAGAAAGCACTACATCTGTAAGCACTGTGTGGGCGAGCCCGAG 120
DB 3065 CTGGCGCCCGCCGAGAAAGTGTCTACATGATGAACACCTGTGTGGGCGAGCCCGAG 3124
QY 121 CTGGAAGGCTTGGCCCTGGAACCCCGGCTGTGGAAGCCGCGAGGGCTGCAAGAGATC 180
DB 3125 CTGGAAGATTTGGCCCTGGAACCCCGGCTGTGGAAGCCGCGAGGGCTGCAAGAGATC 3184
QY 181 ATGAACAGCTGAGCCCGCCCTGAGACCGGCAAGAGAGCTGGCAAGCTGTACAC 240
DB 3185 ATCCGCACTGACCCCGCCCTGAGACCGGCAAGAGAGCTGTGAAGAGCTGTACAC 3244
QY 241 ACCGTGGCCACCTGTGATCTGCTGACGCGCGCATGAGTCTGGCAACCAAGAGGCC 300
DB 3245 ACCGTGGCCACCTGTGATCTGCTGACGCGCGCATGAGTCTGGCAACCAAGAGGCC 3304
QY 301 CTGGAACAGTGGAGAGAGAGAGCAAGTCCCAAGAGAGCCAGAGGCGCAAGAG 360
DB 3305 CTGGAACAGTGGAGAGAGAGAGCAAGTCCCAAGAGAGCCAGAGGCGCAAGAG 3364
QY 361 GCGGAC---GGCAAGGTGAGCCAGAACTACCCCATGTGCAAGAACTGCAAGGCGCAGATG 417
DB 3365 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTGCAAGAACTGCAAGGCGCAGATG 3424
QY 418 GTGACACAGGCCATCAGCCCGCCGACCTGTGAACGCTGTGGTGAAGTGAAGAGAG 477
DB 3425 GTGACACAGGCCATCAGCCCGCCGACCTGTGAACGCTGTGGTGAAGTGAAGAGAG 3484
QY 478 GCGTTAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGAGGAGGCGCACCCCGCAG 537
DB 3485 GCGTTAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGAGGAGGCGCACCCCGCAG 3544
QY 538 GACTTGAACAGATGTTGAACAACCGTGGGCGGCGCACAGGCGCGCATGAGATGCTGAAG 597
DB 3545 GACTTGAACAGATGTTGAACAACCGTGGGCGGCGCACAGGCGCGCATGAGATGCTGAAG 3604
QY 598 GACACCATCAAGAGAGAGGCGCGGAGTGGAGACCGCTGCAACCCGTGAGAGCGGCGCC 657
DB 3605 GACACCATCAAGAGAGAGGCGCGGAGTGGAGACCGCTGCAACCCGTGAGAGCGGCGCC 3664
QY 658 GTGGCGCCCGGAGATGCGGAGACCCCGGCGGAGAGATGCGCGGCGCGCACAGCAGC 717
DB 3665 ATGCGCCCGGAGATGCGGAGACCCCGGCGGAGAGATGCGCGGCGCGCACAGCAGC 3724
QY 718 CTGCAAGAGCAAGTGGCTGTGATGACCAAGCAACCCCGCTGTGGGCGCAACTTAC 777
DB 3725 CTGCAAGAGCAAGTGGCTGTGATGACCAAGCAACCCCGCTGTGGGCGCAACTTAC 3784

```

```

QY 778 AAGCGGTGATGATCTCTGGGCTTGAACAAGATCTGCGGATGTACAGCCCGGTGAGCATC 837
DB 3785 AAGCGGTGATGATCTCTGGGCTTGAACAAGATCTGCGGATGTACAGCCCGGTGAGCATC 3844
QY 838 CTGACATCGGCGAGGAGCCCGCAAGAGGCGCTTCCGGAATACGTGAGACCGCTTCTTCAAG 897
DB 3845 CTGACATCAAGAGAGGAGCCCGCAAGAGGCGCTTCCGGAATACGTGAGACCGCTTCTTCAAG 3904
QY 898 ACCCTGGCGCGGAGAGGAGCCAGAGAGCGTGAAGAACTGATGACCGAGACCTGTCTG 957
DB 3905 ACCCTGGCGCGGAGAGAGGAGCCAGAGAGCGTGAAGAACTGATGACCGAGACCTGTCTG 3964
QY 958 GTGCAAGACGCCAACCCCGCATGCAAGACATCTCTGCGGCTGTGGCCCGGCGCACCC 1017
DB 3965 GTGCAAGACGCCAACCCCGCATGCAAGACATCTCTGCGGCTGTGGCCCGGCGCACCC 4024
QY 1018 CTGGAAGAGATGATGACCGCTGCGAGAGGCGTGGAGGCGCGCGCACAGGCGCGCGTG 1077
DB 4025 CTGGAAGAGATGATGACCGCTGCGAGAGGCGTGGAGGCGCGCGCACAGGCGCGCGTG 4084
QY 1078 CTGGCGAGGCGATGAGCGGCAAGCGTGAACATCATGATCAGAGAGGCGCATCCGCAAG 1137
DB 4085 CTGGCGAGGCGATGAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAG 4141
QY 1138 AAGGCGCCCGGCGCAAGCTCAAGTCTTCAACTGCGGCAAGAGAGGCGCATCCGCAAG 1197
DB 4142 AAGGCGCCCGGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGAGGCGCATCCGCGC 4201
QY 1198 AACTGCGGCGCGCGCGCAAGAGAGGCGTGGAGAGTGGAGAGAGGCGCGCAAGATG 1257
DB 4202 AACTGCGGCGCGCGCGCAAGAGAGGCGTGGAGAGTGGAGAGAGGCGCGCAAGATG 4261
QY 1258 AAGCATGCAAGGAGCGGCGCAAGCGGCAAGCTTCTGCGCAAGATCTGCGGCAAGGAGC 1317
DB 4262 AAGCATGCAAGGAGCGGCGCAAGCGGCAAGCTTCTGCGCAAGATCTGCGGCAAGGAGC 4321
QY 1318 CGCGCGCGCAATTTCTGCAAGAACCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGC 1377
DB 4322 CGCGCGCGCAATTTCTGCAAGAACCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGC 4363
QY 1378 CGCGCGCGGAGAGCTTCTGAGAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGC 1437
DB 4364 CGCGCGCGGAGAGCTTCTGAGAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGC 4423
QY 1438 GACCGCGAGCCCTTACCGGAGGCGCGCTGACCGGCTGCGGAGAGCTGTTGGCGAGCGGCC 1497
DB 4424 G-----ACCGGAGAGCCCTGACAGGCTGGAAGAGCTGTTGGCGAGAGAGCGCC 4471
QY 1498 CTGAGCCAGTAA 1509
DB 4472 CTGAGCCAAATA 4483

```

#### RESULT 14

US-10-190-435-34

/ Sequence 34, Application US/10190435

/ Publication No. US20030143248A1

/ GENERAL INFORMATION:

/ APPLICANT: ZUR MEHDE, Jan

/ APPLICANT: BARNETT, Susan W.

/ APPLICANT: LIAN, Ying

/ APPLICANT: ENGELBRECHT, Susan

/ APPLICANT: VAN RENSBURG, Estrelita J.

/ TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

/ FILE REFERENCE: P18133.003 / 2302-18133

/ CURRENT APPLICATION NUMBER: US/10/190,435

/ NUMBER OF SEQ ID NOS: 319

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 34

/ LENGTH: 4606

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv1-gagmod.BW965
US-10-190-435-34

```

Query Match 84.5%; Score 1275.2; DB 15; Length 4606;  
 Best Local Similarity 92.5%; Pred. No. 9.6e-267;  
 Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

QY 1 ATGGGCGCGCGCGCGCGCATCTGCGCGCGGCGAGAACGTGGAACAAGTGGAGAGATCCGC 60
DB 3128 ATGGGCGCGCGCGCGCGCATCTGCGCGCGGCGCAAGCTGGAACGCTGGAGGCGCATCCGC 3187
QY 61 CTGGCGCGCGCGCGCGCGCAAGACCTAATGCTGGAACCACTGTTGGGCGCGCGCGAG 120
DB 3188 CTGGCGCGCGCGCGCGCAAGAGTGTACATGATGAACACCTGTGTGGCGCGCGCGAG 3247
QY 121 CTGGAAGGCGCTGCGCGCTGGAACCGCGCGCTGTGGAAGCCGCGAGGCGTGAAGCAGATC 180
DB 3248 CTGGAAGAACTTCCGCTTGAACCCCGCGCTGTGGAAGCCAGCGAGGCTGCAAGCAGATC 3307
QY 181 ATGAAGCAGCTGCAAGCGCGCGCTGCAAGCGCGCAGAGAGCTGCGCAGCTGTACAA 240
DB 3308 ATCCGCGCAGCTGCAAGCGCGCGCTGCAAGCGCGCAGAGAGCTGGAAGAGCTGTTCAG 3367
QY 241 ACCGTGGCCACCTCTGTAATGCTGCGTGCAGCGCGCATGAGAGTCCGCGCAACCAAGAGCGC 300
DB 3368 ACCGTGGCCACCTCTGTAATGCTGCGTGCAGAGAGATGAGAGTCCGCGCAACCAAGAGCGC 3427
QY 301 CTGGAACAAGTGAAGGAGAGAGCAACAAGTCCAGCAGAGAGCCAGAGGCGCAAGAG 360
DB 3428 CTGGAACAAGTGAAGGAGAGAGCAACAAGTCCAGCAGAGAGCCAGAGGCGCAAGAG 3487
QY 361 GCGGAC--GCGAAGGTGAAGCAAACTACCCCATGTGCAAGAACTGAGGCGCGCAGATG 417
DB 3488 GCGGACAAAGGCAAGGTGAAGCAAACTACCCCATGTGCAAGAACTGAGGCGCGCAGATG 3547
QY 418 GTGCAACCAAGCCATCAAGCCCGCGCAACCGCTGAGTGAAGTATCCAGAGAGAG 477
DB 3548 GTGCAACCAAGCCATCAAGCCCGCGCAACCGCTGAGTGAAGTATCCAGAGAGAG 3607
QY 478 GCGTTAGCGCGCGAGGTGATCCCATGTTCAACGCGCTGAGAGAGGCGCGCAACCCCGCAG 537
DB 3608 GCGTTAGCGCGCGAGGTGATCCCATGTTCAACGCGCTGAGAGAGGCGCGCAACCCCGCAG 3667
QY 538 GACCTGAACAGATGTTGAACAACCGTGGCGCGCACAGGCGCGCATGCAATGCTGAAG 597
DB 3668 GACCTGAACAGATGTTGAACAACCGTGGCGCGCACAGGCGCGCATGCAATGCTGAAG 3727
QY 598 GACACCATCAAGAGAGAGCGCGCGAGTGGAGACCGGCTGCAACCCCGTGAAGCGCGCGCC 657
DB 3728 GACACCATCAAGAGAGAGCGCGCGAGTGGAGACCGGCTGCAACCCCGTGAAGCGCGCGCC 3787
QY 658 GTGGCGCGCGCGCGAGATGCGCGCAACCGCGCGGAGAGATGCGCGCGCGCAACAGCACC 717
DB 3788 ATCGCGCGCGCGCGAGATGCGCGCAACCGCGCGGAGAGATGCGCGCGCGCAACAGCACC 3847
QY 718 CTGCAAGAGCAAGTGCCTGTGATGACAGCAACCCCGCGTGGCGCGAGCATCTAC 777
DB 3848 CTGCAAGAGCAAGTGCCTGTGATGACAGCAACCCCGCATCGTGGCGCGAGCATCTAC 3907
QY 778 AAGCGGTGATCATCTCTGGGCGCTGAACAAGATGTGCGAGTGAACGCGCGTGAAGATC 837
DB 3908 AAGCGGTGATCATCTCTGGGCGCTGAACAAGATGTGCGAGTGAACGCGCGTGAAGATC 3967
QY 838 CTGGAACATCGCGCGAGGCGCGCAAGAGAGCCCTTCGCGCACTAGCTGAGACGCTTCTTCAAG 897
DB 3968 CTGGAACATCGCGCGAGGCGCGCAAGAGAGCCCTTCGCGCACTAGCTGAGACGCTTCTTCAAG 4027
QY 898 ACCCTGCGCGCGAGAGGCGCAACCGAGAGCTGAAGAACTGATGACCGAGAGCCCTGCTG 957

```

```

DB 4028 ACCCTGCGCGCGAGAGAGAGCAACCGAGAGGTGAAGAACTGATGACCGAGACCCCTGCTG 4087
QY 958 GTGCAAGAACCGCAACCCCGCATCTGCAAGACCAATCTCTGCGCGCTCTCGGCGCGCGCAC 1017
DB 4088 GTGCAAGAACCGCAACCCCGCATCTGCAAGACCAATCTCTGCGCGCTCTCGGCGCGCGCAC 4147
QY 1018 CTGGAAGAGATGATGACCGCGCTGCCAGAGGCGCTGGGCGCGCGCGCGCAACAAGCGCGCTG 1077
DB 4148 CTGGAAGAGATGATGACCGCGCTGCCAGAGGCGCTGGGCGCGCGCGCGCAACAAGCGCGCTG 4207
QY 1078 CTGGCGCGAGCGATGAGCCAGCGCAACAGCTGAAATCATGATGACAGAGCAACTTC 1137
DB 4208 CTGGCGCGAGCGATGAGCCAGCGCAACAGCTGAAATCATGATGACAGAGCAACTTC 4264
QY 1138 AAGGCGCGCGCGCGCAAGTCAAGTCTTCAACTGCGCGCAAGAGGCGCGCATGCGCAAG 1197
DB 4265 AAGGCGCGCGCGCGCATGTCAAGTCTTCAACTGCGCGCAAGAGGCGCGCATGCGCGCAAG 4324
QY 1138 AACTGCGCGCGCGCGCGCAAGAGGCGCTGTGAAAGTCCGCGCAAGAGGCGCGCATGAG 1257
DB 4325 AACTGCGCGCGCGCGCGCAAGAGGCGCTGTGAAAGTCCGCGCAAGAGGCGCGCATGAG 4384
QY 1258 AAGACTGACCGAGCGCGCAAGGCGCAACTTCTGTGGCAAGATCTGCGCAAGAGGCGC 1317
DB 4385 AAGACTGACCGAGCGCGCAAGGCGCAACTTCTGTGGCAAGATCTGCGCAAGAGGCGC 4444
QY 1318 CGCGCGCGCAACTCTCTGCAAGAACCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1377
DB 4445 CGCGCGCGCAACTCTCTGCAAGAGCGG-----CGCGAGCGCGCGCGCGCGCGCGCGCGCG 4486
QY 1378 CGCGCGCGCGAGCGCTTCCGCTTGAAGAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGAG 1437
DB 4487 CGCGCGCGCGAGCGCTTCCGCTTGAAGAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGAG 4546
QY 1438 GACCGGAGCGCTTACCGCGAGCGCGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
DB 4547 G-----ACCGGAGAGCGCTTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4594
QY 1498 CTGAGCCAGTAA 1509
DB 4595 CTGAGCCAAATA 4606

```

```

RESULT 15
US-10-190-435-36
; Sequence 36, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEERDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGBLERT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 4615
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv2-gagmod.BW965
US-10-190-435-36

```

Query Match 84.5%; Score 1275.2; DB 15; Length 4615;  
 Best Local Similarity 92.5%; Pred. No. 9.6e-267;  
 Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGCGCCGCGCAGCATCTGCGCGGCGAGAGTGGGGAAGATCCCG 60  
 Db 3137 ATGGGCGCCGCGCAGCATCTGCGCGGCGGAGAGTGGGGAAGATCCCG 3196  
 QY 61 CTGCGCGCCGCGCGCAAGAGTCTACATGCTGAGCACTGTGTGGCCAGCCGAG 120  
 Db 3197 CTGCGCGCCGCGCGCAAGAGTCTACATGATGAACACCTGGTGTGGGCGAGCCGAG 3256  
 QY 121 CTGGAAGGCTTGGCCCTGGAACCCCGGCTGTGGAACCGCGAGGGCTGACAGCATC 180  
 Db 3257 CTGGAAGAGTGGCCCTGGAACCCCGGCTGTGGAACCGCGAGGGCTGCAAGCATC 3316  
 QY 181 ATGAAGCAGCTGAGCGCCGCTGAGACCGGACCGAGAGGTGGCGACCTGTACAC 240  
 Db 3317 ATCCGCGACCTGACCCCGCTGAGACCGGAGAGAGAGCTTGAAGAGCTGTTCAC 3376  
 QY 241 ACCGTGGCACCCTGTACTGTGCGGTGACGCGCGGATGAGGTCCGACACCAAGAGGCC 300  
 Db 3377 ACCGTGGCACCCTGTACTGTGCGGTGACGAGAAAGATGAGGTCCGACACCAAGAGGCC 3436  
 QY 301 CTGGAACAAGATGAGAGAGAGAGCAAGTCCACAGAAAGCCAGCAGGCCAAGAG 360  
 Db 3437 CTGGAACAAGATGAGAGAGAGAGCAAGTCCAGCAGAAAGTCCAGCAGGCCCGAGGCC 3496  
 QY 361 GCCGAC---GGCAAGGTGAGCCGAACTACCCCATGTGCAAGACCTGAGAGGCCGAGATG 417  
 Db 3497 GCCGACAAAGGCAAGGTGAGCCGAACTACCCCATGTGCAAGACCTGAGAGGCCCGAGATG 3556  
 QY 418 GTGCAACAGGCCATCAAGCCCGGCACTGAAAGCTGTGGTGAAGGTGATCGAGAGAA 477  
 Db 3557 GTGCAACAGGCCATCAAGCCCGGCACTGAAAGCTGTGGTGAAGGTGATCGAGAGAA 3616  
 QY 478 GCTTCAAGCCCGGAGGTGATCCCATGTTCAACCGCTTGAAGAGAGGCCACCCCGAG 537  
 Db 3617 GCTTCAAGCCCGGAGGTGATCCCATGTTCAACCGCTTGAAGAGAGGCCACCCCGAG 3676  
 QY 538 GACCTGAACAGCATGTGTAACAACCGTGGGCGGCAACAGCGGCCATGCAATGCTGAAG 597  
 Db 3677 GACCTGAACAGCATGTGTAACAACCGTGGGCGGCAACAGCGGCCATGCAATGCTGAAG 3736  
 QY 598 GACACCATCAACGAGAGAGCGCGGAGTGGGACCGGCTGACCCCTGCAAGGCCCGGCC 657  
 Db 3737 GACACCATCAACGAGAGAGCGCGGAGTGGGACCGGCTGACCCCGTGCACCGCGGCC 3796  
 QY 658 GTGGCCCCCGGCGCATGTGCGCAACCCCGCGGCGAGCATGCGCGGCCACAGCAC 717  
 Db 3797 ATCGCCCCCGGCGCATGTGCGCAACCCCGCGGCGAGCATGCGCGGCCACAGCAC 3856  
 QY 718 CTGCAAGACAGATGCGCTGATGACAGCAACCCCGGCGGCGGCGCATCTAC 777  
 Db 3857 CTGCAAGACAGATGCGCTGATGATGACAGCAACCCCGCATCCCGTGGCGCATCTAC 3916  
 QY 778 AAGCGGTGATCATCTGGGCTGAAACAAGATGATGCGATGATACAGCCCGTGAAGCATC 837  
 Db 3917 AAGCGGTGATCATCTGGGCTGAAACAAGATGATGCGATGATACAGCCCGTGAAGCATC 3976  
 QY 838 CTGACATCCGCAAGGCGCCCAAGAGCCCTTCCGCGCATGCTGACCGCTTCTTCAAG 897  
 Db 3977 CTGACATCCGCAAGGCGCCCAAGAGCCCTTCCGCGCATGCTGACCGCTTCTTCAAG 4036  
 QY 898 ACCCTGGCGCGGAGAGGCCCAACGAGCGTGAAGAACTGATGACCGAGACCTGTG 957  
 Db 4037 ACCCTGGCGCGGAGAGGCCCAACGAGCGTGAAGAACTGATGACCGAGACCTGTG 4096  
 QY 958 GTGCAAGACGCAACCCCGCATGCAAGACATCTGCGGCTCTGCGGCCCGGCGCAC 1017  
 Db 4097 GTGCAAGACGCAACCCCGCATGCAAGACATCTGCGGCTCTGCGGCCCGGCGCAC 4156  
 QY 1018 CTGGAAGAGATGATGACCGCTGCGAGGCGTGGGCGGCCCGGCAAGGCCCGCGTG 1077  
 Db 4157 CTGGAAGAGATGATGACCGCTGCGAGGCGTGGGCGGCCCGGCAAGGCCCGCGTG 4216  
 QY 1078 CTGGCCGAGGCGATGAGCGAGCAACAGCTGGAACATGATGACGAAGCAACTTC 1137

Db 4217 CTGGCCGAGGCGATGAGCCAGGCCAAACAC---ACGTGATGATCAGAGAGCAACTTC 4273  
 QY 1138 AAGGCGCCCGGCGCAAGCTCAAGTGTCTTCAACTGCGGCAAGAGAGGCCACATGCGCAAG 1197  
 Db 4274 AAGGCGCCCGGCGCATGCTTCAAGTGTCTTCAACTGCGGCAAGAGAGGCCACATGCGCGC 4333  
 QY 1198 AACTGCGGCGCCCGGCAAGAGGCTGTGGAAGTGTGGGCAAGAGAGGCCACCAAGATG 1257  
 Db 4334 AACTGCGGCGCCCGGCAAGAGGCTGTGGAAGTGTGGGCAAGAGAGGCCACCAAGATG 4393  
 QY 1258 AAGGACGACCGAGCGCCAGGCGCAACTTCTGGGCAAGATGTGGCCCGCAGCAAGAGGC 1317  
 Db 4394 AAGGACGACCGAGCGCCAGGCGCAACTTCTGGGCAAGATGTGGCCCGCAGCAAGAGGC 4453  
 QY 1318 CGCCCGGCAACTTCTGTGAGAACCGCAGCGAGCCCGCGCCCAAGCTGTGCCACCGCC 1377  
 Db 4454 CGCCCGGCAACTTCTGTGAGAGCGG-----CGCGAGGCCACCGCC 4495  
 QY 1378 CCCCCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCCCAAGAGAGGCCCAAG 1437  
 Db 4496 CCCCCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCGAGAGAGAGCAAG 4555  
 QY 1438 GACCGGAGCCCTACCGGAGGCCCTGAACCGGCTGGCGAGCTGTGGGCGAGGGCCCC 1497  
 Db 4556 G-----ACCGGAGAGCCTTGAACAGCTGAAGAGCTGTGGGCGAGGCC 4603  
 QY 1498 CTGAGCCAGTAA 1509  
 Db 4604 CTGAGCCAAATA 4615

Search completed: September 1, 2005, 23:17:25  
 Job time : 943.802 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 1, 2005, 21:17:23 ; Search time 26.261 Seconds  
(without alignments)  
8578.900 Million cell updates/sec

Title: US-09-475-704A-4

Perfect score: 2840  
Sequence: 1 atggcgccgcgcgcagcat.....ggcgccctgagccagtaa 1509

Scoring table:

BLOSUM62		
Xgapop 10.0 , Xgapext 0.5		
Ygapop 10.0 , Ygapext 0.5		
Fgapop 6.0 , Fgapext 7.0		
Delop 6.0 , Delext 7.0		

Searched: 513545 segs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=x1p  
-O=/cgn2\_1/USPTO.spool.p/US09475704/rnat.01092005.160751.19541/app\_query.fasta\_1.3342  
-DB=Issued Patents\_AA -QFMT=fastaan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humand0.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORExt -HSPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09475704@cgn\_1.53 @runat.01092005.160751.19541 -NCPQ=6 -ICPU=3  
-NO WMAP -LANGBOUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	86.5	492	4	US-09-991-258-5
2	2246	79.1	500	1	US-08-375-510-1
3	2246	79.1	500	2	US-08-487-657-1
4	2246	79.1	500	4	US-09-309-572-16
5	2246	79.1	500	4	US-09-718-096-16
6	2238	78.8	1350	4	US-09-952-060-35
7	2237	78.8	512	3	US-08-463-210-8
8	2237	78.8	512	3	US-09-124-900-2
9	2237	78.8	512	4	US-08-463-028-8
10	2223	78.3	500	2	US-08-816-155B-45
11	2223	78.3	500	3	US-09-079-587-45
12	2094	73.7	493	4	US-09-952-060-33

13	2018.5	71.1	478	1	US-08-127-499A-11	Sequence 11, Appl
14	2018.5	70.8	478	1	US-08-482-847-11	Sequence 11, Appl
15	2011	70.8	437	3	US-08-392-794A-2	Sequence 2, Appl1
16	1949	68.6	512	4	US-09-319-588C-4	Sequence 4, Appl1
17	1748.5	61.6	498	1	US-08-470-202-59	Sequence 59, Appl
18	1748.5	61.6	498	1	US-08-471-770-59	Sequence 59, Appl
19	1748.5	61.6	498	2	US-08-468-059-59	Sequence 59, Appl
20	1748.5	61.6	498	3	US-09-109-916-59	Sequence 59, Appl
21	1748.5	61.6	498	3	US-09-886-156-59	Sequence 59, Appl
22	1748.5	61.6	498	4	US-09-886-149-59	Sequence 59, Appl
23	1748.5	61.6	498	4	US-09-886-150-59	Sequence 59, Appl
24	1748.5	61.6	498	4	US-09-886-159-59	Sequence 59, Appl
25	1748.5	61.6	498	4	US-10-325-090-59	Sequence 59, Appl
26	1726.5	60.8	498	1	US-08-470-202-60	Sequence 60, Appl
27	1726.5	60.8	498	1	US-08-471-770-60	Sequence 60, Appl
28	1726.5	60.8	498	2	US-08-468-059-60	Sequence 60, Appl
29	1726.5	60.8	498	3	US-09-109-916-60	Sequence 60, Appl
30	1726.5	60.8	498	4	US-09-886-156-60	Sequence 60, Appl
31	1726.5	60.8	498	4	US-09-886-149-60	Sequence 60, Appl
32	1726.5	60.8	498	4	US-09-886-150-60	Sequence 60, Appl
33	1726.5	60.8	498	4	US-09-886-159-60	Sequence 60, Appl
34	1726.5	60.8	498	4	US-10-325-090-60	Sequence 60, Appl
35	1671.5	58.9	363	2	US-08-850-049-130	Sequence 130, App
36	1671.5	58.9	363	2	US-08-050-478-130	Sequence 130, App
37	1671.5	58.9	363	3	US-09-414-117-130	Sequence 130, App
38	1671.5	58.9	363	4	US-09-678-437-130	Sequence 130, App
39	1671.5	58.9	363	4	US-09-943-722-130	Sequence 130, App
40	1545.5	54.4	799	1	US-07-648-796A-7	Sequence 7, Appl1
41	1544	54.4	458	1	US-07-648-796A-5	Sequence 5, Appl1
42	1530.5	53.9	337	1	US-07-648-796A-1	Sequence 1, Appl1
43	1514.5	53.3	328	4	US-08-776-188C-77	Sequence 77, Appl
44	1453.5	51.2	518	4	US-09-206-551-45	Sequence 45, Appl
45	1410.5	49.7	294	3	US-09-370-368-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-991-258-5  
; Sequence 5, Application US/09991258  
; Patent No. 6783939  
; GENERAL INFORMATION:  
; APPLICANT: Olmsted, Robert  
; APPLICANT: Kelch, Paula  
; APPLICANT: Dryga, Sergey  
; APPLICANT: Caley, Ian  
; APPLICANT: Maughan, Maureen  
; APPLICANT: Johnston, Robert  
; APPLICANT: Davis, Nancy  
; APPLICANT: Swanson, Ronald  
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE  
; FILE REFERENCE: 0113.0001U3  
; CURRENT APPLICATION NUMBER: US/09/991,258  
; CURRENT FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 09/902,537  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/216,995  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. 6783939 =  
; OTHER INFORMATION: synthetic construct  
US-09-991-258-5  
Alignment Scores:  
Pred. No.: 2,32e-141 Length: 492  
Score: 2457.00 Matches: 465

Percent Similarity: 95.22% Conservative: 13  
 Best Local Similarity: 92.63% Mismatches: 14  
 Query Match: 86.51% Indels: 10  
 DB: 4 Gaps: 2

US-09-475-704a-4 (1-1509) x US-09-991-258-5 (1-492)

```

QY 1 ATGGGCGCCGCGCAGCATCTCTGCGGCGAGAGAGCTGAGCAAGTGGAGAGATCCGC 60
DB 1 MetAlaAlaArgAlaSerIleLeuArgGlyGluLeuAspArgTrpGluYsIleArg 20
QY 61 CTGGCGCCGCGCGCAGAGAGCACTACATGCTGAGACACCTGTGTGGCCAGCCGCGAG 120
DB 21 LeuArgProGlyGlyYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYs 40
QY 121 CTGGAGGGCTTGGCCCTGAAACCCCGGCTGTGGAGAACCCGCGAGGCTGGCAAGCATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysYsGlnIle 60
QY 181 ATGAAGCAGCTGAGCGCCGCTGACAGCCGCGACCGAGAGAGCTGCGACCTGTACAC 240
DB 61 MetYsGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuYsSerLeuYsAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGTCACGCCGCGCATGAGGTCCGCGACCAAGAGGCC 300
DB 81 ThrValAlaThrLeuYsCysValHisGlyYsIleGluValArgAspThrYsGluAla 100
QY 301 CTGGACAAGATGAGAGAGAGAGAGAGCAAGTCCACAGAGAGAGAGAGAGAGAGAG 360
DB 101 LeuAspYsYsIleGluGluGluGlnAsnYsCysGlnGlnYsThrGlnIleAlaYsAla 120
QY 361 GCCGACGCGAAGGTGAGCAGAGACTACCCCATGCTGACAGACCTGAGAGAGAGATGATG 420
DB 121 AlaAspGlyYsValSerGlnAsnYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYs 140
QY 421 CACGAGGCGCATGAGCCCGCGACCTGAAAGCCCTGGGTGAAGGTATGAGAGAGGCC 480
DB 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValYsValIleGluGluYsAla 160
QY 481 TTCAGCGCCGAGGTATGCCCATGTTTCAGCGCCCTGAGAGAGAGAGAGAGAGAGAG 540
DB 161 PheSerProGluValIleProMetPheThrAlaLeuSerGlnYsAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTTGAACACCGTGGCGGCGCAGAGCCGCGCATGAGATGCTGAAGAC 600
DB 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMetLeuYsAsp 200
QY 601 ACCATCAACGAGAGAGCGCCGAGTGGACCGCTGACACCCGCTGACAGGCCGCGCCG 660
DB 201 ThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisProValHisIleGlyProIle 220
QY 661 GCGCCGCGGAGATGGCGGAGCCCGCGGCGAGAGATGCGCGGCGCGCAGAGACCTTG 720
DB 221 AlaProGlyGlnMetArgGluProAlaGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGCAGATGCGCTGATGACAGAGAACCCCGCTGCGCGTGGCGGCGACATTTACAAG 780
DB 241 GlnGluGlnIleAlaIleTrpMetThrSerAsnProProIleProValGlyAspIleYsTrp 260
QY 781 CCGTGATATCTCTGGGCGCTGAACAAGATGTCGGAGATGATACAGCCCGCTGAGATCTTG 840
DB 261 ArgTrpIleIleLeuGlyLeuAsnYsIleValArgMetYsSerProValSerIleLeu 280
QY 841 GACATCCGCGCAGGCGCCCAAGAGCCCTCCGCGACTAGTGAACCGCTTTCTTACAAGAC 900
DB 281 AspIleArgGlnGlyProYsGluProPheArgAspTrpValAspArgPheLeuYsThr 300
QY 901 CTGGCGCGCGCAGAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 301 LeuArgIleGluGlnIleAlaThrGlnGluValYsAsnTrpMetThrAspThrLeuVal 320
QY 961 CAGAAAGCCCAAGCCGAGCTGCAAGACCATCTCGCGGCTCTCGCGCCGCGCGCACCTTG 1020

```

```

DB 321 GlnAlaAlaAsnProAspCysLeuYsThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGAGAGATGATACCGCTGCGAGGCGGTGGCGGCGCCGCGCAGAGAGAGAGAGAGAG 1080
DB 341 GlnGluMetMetThrAlaCysGlnGlyAlaGlyGlyProGlyHisYsAlaArgValLeu 360
QY 1081 GCGGAGGAGATGAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 361 AlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPheYs 380
QY 1141 GCGCCCGCGCGCAAGCTGATGCTTCAATGCGGCAAGAGAGAGAGAGAGAGAGAGAG 1200
DB 381 GlyProArgArgIleValIleCysArgPheAsnYsGlyYsGluGlnHisIleAlaArgAsn 400
QY 1201 TGGCGGCGCCCGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 401 CysArgAlaProArgYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYsYs 420
QY 1261 GACTCAGCAGAGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyYsIleTrpProSerHisYsYsYsYs 440
QY 1321 CCGGCAACTTCTGTCAGAACCGCGAGCGAGCCGCGCCGCGCCGCGCCGCGCCGCG 1380
DB 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454
QY 1381 CCGCGCGAGAGCTTCCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProYsGlnGluProIleGlu 474
QY 1441 CGCGAGCCCTTACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 475 -----ArgGluProLeuThrSerLeuYsSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
DB 491 SerGln 492

```

RESULT 2  
 US-08-375-510-1  
 / Sequence 1, Application US/08375510  
 / Patent No. 5576421  
 / GENERAL INFORMATION:  
 / APPLICANT: Saito, Atsushi  
 / APPLICANT: Sinagawa, Hideo  
 / APPLICANT: Nakata, Atsuo  
 / TITLE OF INVENTION: HIV ANTIGEN  
 / NUMBER OF SEQUENCES: 4  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Birch, Stewart, Kolasch and Birch  
 / STREET: P. O. Box 747  
 / CITY: Falls Church  
 / STATE: Virginia  
 / COUNTRY: U.S.A.  
 / ZIP: 22040-0747  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / OPERATING SYSTEM: IBM PC compatible  
 / SOFTWARE: Patentin Release #1.0, Version #1.25  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/375,510  
 / FILING DATE:  
 / CLASSIFICATION: 435  
 / PRIORITY APPLICATION DATA:  
 / APPLICATION NUMBER: US/07/985,949  
 / FILING DATE: 04-DEC-1992  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Svensson, Leonard R.  
 / REGISTRATION NUMBER: 30,330  
 / REFERENCE/DOCKET NUMBER: 216-309P  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (703) 241-1300

TELEFAX: (703) 241-2848  
 TELEX: 248345  
 INFORMATION FOR SRO ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 500 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Human immunodeficiency virus type 1  
 US-08-375-510-1

## Alignment Scores:

Prod. No.:	1,42e-128	Length:	500
Score:	2246.00	Matches:	425
Percent Similarity:	90.16%	Conservative:	33
Best Local Similarity:	83.66%	Mismatches:	36
Query Match:	79.08%	Indels:	14
DB:	1	Gaps:	5

US-09-475-704a-4 (1-1509) x US-08-375-510-1 (1-500)

```

QY 1 ATGGGGGCGGCGCGCATCTGCGGCGGAGAACATGGAAGATCCG 60
   1 MeGtYAlaIarYgAlaSerGlyGlyLeuAspYstrpGlyStrIleArY 20
QY 61 CTGCGCGCGCGGCGGAGAAAGCACTACATCTGAGCACTGTGTGGCGCGCGAG 120
   21 LeuArgProGlyGlyIySylSgIntYrYlSleuYhIstIleValTrpAlaSerArGln 40
QY 121 CTGAGAGGCTTGCGCTGAAACCCGCGCTGTGAGACCGCGAGGCTGTGACAGATC 180
   41 LeuGluArgPheAlaValaIaenProGlyLeuLeuGluThrSerGlyCysArgGlnIle 60
QY 181 ATGAGACAGCTGACGCGCGCGCGCGAGACCGGACGAGAGTGGCGGCGCTGTACAC 240
   61 LeuGlyGlnIleuInProSerLeuGlnIntGlySerGlnIleuArgSerLeuYrAan 80
QY 241 ACCGTGGCGACCTGTACTGCTGTGACGCGCGCATGAGAGTCCGAGACCAAGAGGCG 300
   81 ThrIleAlaValLeuYrCysValaHISGlnArgIleAspValIyAspThrIySgIua 100
QY 301 CTGAGCAAGATCGAGAGAGAGCAAGAAAGTCCACAGACAGACCCAGAGCGCAAGAG 360
   101 LeuAspIyIleGlnIleuGlnIleuAsnYserIySylValaGlnIleAlaIa 120
QY 361 GCGGACGGC-----AAGTGAACCAAGACTACCCATCGTGCAAACTGTGCGAGGC 411
   121 AspThrGlyAAsnAsnSerGlnValSerGlnAsnTrpProIleValGlnAsnIleuGlnIy 140
QY 412 CAGATGTGACACGAGCATGACCGCGCACCTTAAAGCGCGGTGAAGTATCGAG 471
   141 GlnMetValHISGlnIleAlaIleSerProAlaGlnThrLeuAsnAlaIleTrpValIySylValGln 160
QY 472 GAGAAAGCTTCAAGCCCGAGGTATCCCATGTTCACCGCCCTGACGAGGCGCGCAC 531
   161 GluYsAlaPheSerProGluValIlePrometPheSerAlaIleuSerGlnIyAlaThr 180
QY 532 CCCGAGGACCTGAACAGCATGTTGAAACACCGTGGGCGGCGACCGCGCATGACAGATG 591
   181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHISGlnIleAlaMetGlnMet 200
QY 592 CTGAAGAGACCATCAAGAGAGCGCGCGAGTGGAGCCGCTGACCCCGTGAAGCGGC 651
   201 LeuYsGlnThrIleAsnGlnIleuAlaIleGlnIleuAspArgIleuHISProValHISAla 220
QY 652 GGGCGCGTGGCGCGCGCGAGATGCGGACCCCGCGGAGAGACATGCGCGCGCGCAC 711
   221 GlyProIleAlaProGlyGlnMetArgIleuProArgIySerAspIleAlaGlyThrThr 240

```

```

QY 712 AGCACCCTGAGAGAGAGATGCGCTGTGATGACCAAGACCCCGCGCTGCGGCGCAC 771
   241 SerThrLeuGlnIleuGlnIleGlyTrpMetThrHISAsnProIleProValGlyGln 260
QY 772 ATCTCAAGCGGTGATCATCTCGGCGCGTGAACAAGATGTGCGGATGTGACGCCCGG 831
   261 IleTrpYsArgTrpIleIleLeuGlyLeuAsnIyIleValArgMetIySerProThr 280
QY 832 AGCATCTTGACATCGCGCGAGCGCCCAAGAGACCTTCCGACTACGTGACCGCTTC 891
   281 SerIleLeuAspIleArgGlnIyProIySgIuProPheArgAspTrpValaIaPargPhe 300
QY 892 TTCGAAGACCTTCGCGCGCGAGCGACCGCGACCGGTGAAGATGTGATGACCGAGAC 951
   301 TrpYsThrLeuArgAlaGlnIleuAlaSerGlnIleuValYsAsnTrpMetThrGlnThr 320
QY 952 CTGCTGTGTGAGAAAGCGCAACCCCGAGCTGCAAGACATCTGCGGCGTTCGCGCGCG 1011
   321 LeuLeuValGlnAsnAlaAsnProAspCysIyStrIleLeuYsAlaIleuGlyProGly 340
QY 1012 GCCACCTGTGAGAGATGATGACCGCTGCGCGAGCGCGCGCGCGCGCGCACAGGCC 1071
   341 AlaThrLeuGlnIleuMetThrAlaCysGlnIyValGlyProGlyTrpIySylAla 360
QY 1072 CGCGTGTGTGCGGAGCGCATGACCGAG--GCCAACAGCGTGAACATCATGATGACGAAG 1128
   361 ArgValaIleuAlaGlnIleuAlaMetSerGlnValThrAsnProAlaThrIleMetIleGlnYs 380
QY 1129 AGCACTTCAGAGGCGCGCGCGCGCAAGCTGACAGTCTTCATCTGCGGCGAGAGGCGCAC 1188
   381 GlyAsnPheArgAsnGlnIleuArgIyStrValIySylPheAsnCysIySylGlnIyHIS 400
QY 1189 ATGCGCAAGACTGTGCGCGCGCGCGCGCAAGAGGCGCTGTGAGATGTGCGGAGAGGCG 1248
   401 IleAlaYsAsnCysArgAlaProArgIySylSgIyStrIySylCysGlyIySylGlnIy 420
QY 1249 CACGAGATGAAGATGTGACCGAGCGCGCGCGCAACTTCTGTGGGCAAGATGTGCGCGCAC 1308
   421 HisGlnMetIyAspCysThrIleuArgIleuAlaAsnPheLeuGlyIleTrpProSer 440
QY 1309 CACAGGGCGCGCGCGCGCACTTCTGTGAGAAACCGCACGAGCGCGCGCGCGCGCGCG 1368
   441 HisIySylGlyArgProGlyAsnPheLeuGlnIleuArgProGln----- 454
QY 1369 CCCACCGCGCGCGCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCGCG 1422
   455 ProThrAlaProProGlnIleuSerPheArgPheGlyGlnIleuThrThrProSerGln 474
QY 1423 AAGCGAGAGCCCAAGAGACCGGAGCGCTTACCGGAGCGCGCTGTGACCGCGCTGTGCGAGCTTG 1482
   475 IySgIleuIleuProIleAspIySylIleuYr-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTCGCGAGCGCGCGCGCTGAGCGAG 1506
   493 PheGlySerAspProSerSerGln 500

```

US-08-487-657-1  
 Sequence 1, Application US/08487657  
 Patent No. 5834267

## GENERAL INFORMATION:

APPLICANT: Saito, Atsushi  
 APPLICANT: Sinagawa, Hideo  
 APPLICANT: Nakata, Aetsuo  
 TITLE OF INVENTION: HIV ANTIGEN  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Birch, Stewart, Kolasch and Birch  
 STREET: P. O. Box 747  
 CITY: Falls Church  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 22040-0747



```
Db          493 PheGlySerAspProSerSerGln 500
|||||
RESULT 4
US-09-309-572-16
; Sequence 16, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; EARLIER FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 500
; TYPE: PRF
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: gag polypeptide
US-09-309-572-16

Alignment Scores:
Pred. No.:      1,42e-128      Length:      500
Score:          2246.00        Matches:      425
Percent Similarity: 90.16%      Conservative: 33
Best Local Similarity: 83.66%      Mismatches:  36
Query Match:    79.08%         Indels:       14
DB:              4              Gaps:         5

US-09-475-704a-4 (1-1509) x US-09-309-572-16 (1-500)
QY      1  ATGGGCGCGCGCGCGCATCTGCGCGCGAGAGAGCTGGAGCAAGTGGAGAGATCCGC 60
Db      1  MetGlyAlaArgAlaSerValIleuSerGlyGlyIleuAspLysTrpGluLysIleArg 20
QY      61  CTGGCGCGCGCGCGCGCAAGAGCACTACATGCTGAGACCTGGTGTGGCGCGCGAG 120
Db      21  LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValIleAlaSerArgGln 40
QY      121  CTGGAGGGCTTGGCTGAAACCCGCGCTGCTGGAGAACCGCGAGGGCTGGCAAGCATC 180
Db      41  LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY      181  ATGAAGCAGCTGACGCGCGCGCGCGAGACCGGACCGAGAGCGTGGCGAGCTGTACAC 240
Db      61  LeuGlyGlnIleuGlnProSerLeuGlnIleGlySerGluGlnIleuArgSerLeuTyrAsn 80
QY      241  ACCGTGGCCACCTGTACTGCTGTGACGCGCGCGCATGAGTCCGCGACCAAGAGAGCC 300
Db      81  ThrIleAlaValLeuTyrCysValHisGlnIleArgIleAspValLysAspThrLysGlnAla 100
QY      301  CTGGACAAATGAGAGAGAGAGAGCAAGTCCCGACAGAGAACCCGACGAGCGCAAGAG 360
Db      101  LeuAspLysIleGluGlnIleuGlnIleuAsnLysSerLysLysAlaGlnIleAlaAla 120
QY      361  GCCGAGCGC-----AAGGTGAGCGAGAACTACCCCATGGTGAGCAACCTGCGAGGGC 411
Db      121  AspThrGlyAsnAsnSerGlnValIleSerGlnIleuTyrProIleValGlnIleuGlnGly 140
QY      412  CAGATGTCACACGAGCGCATCAGCGCGCGCATGAGCGCTGGTGAAGGTGATCGAG 471
Db      141  GlnMetValHisGlnIleIleSerProArgThrLeuAsnAlaIlePheValLysValValGln 160
QY      472  GAGGAAGCTTTCAGCGCGCGAGGTGATCCCATATTTACACGCGCTTGAGCGAGGGCGCAC 531
Db      161  GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY      532  CCCGAGGACTGAACAGATGTTGAACACGCTGGCGGCGCACGAGCGCGCATGCGAGT 591
|||||
```

```
Db      181  ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnIleAlaMetGlnMet 200
QY      592  CTGAAGACACCATTAAGAGAGAGCGCGCGAGTGGAGCCGCTTCACCCCGCTGAGAGCC 651
Db      201  LeuLysGluThrIleAsnGluGlnIleAlaIleGluTyrAspArgLeuHisProValHisAla 220
QY      652  GGGCGCGCGCGCGCGAGATGCGCGAGACCGCGCGGAGAGCATGCGCGCGCGCAC 711
Db      221  GlnProIleAlaProGlyGlnMetArgGluProArgLysSerAspIleAlaGlyThrThr 240
QY      712  AGCACCCTGACAGAGACAGATCGCTGTGATGACCAAGCAACCCCGCTGCGCGCGAGC 771
Db      241  SerThrLeuGlnIleGlnIleGlyTyrMetThrHisAsnProProIleProValGlyGln 260
QY      772  ATCTCAAGCGGTGATATCTGTGGCTTGAACAAAGATGCTGGAGTGTACAGCCCGCTG 831
Db      261  IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProThr 280
QY      832  AGCATCTGGAACATCCGCGAGCGCGCGCAAGAGCGCTCCGCGACTAGTGGAGCGCTTC 891
Db      281  SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY      892  TTCAGACCTGTCGCGCGCGAGAGCGCACCGCAGACGTAAGAACTGGATGACCGAGACC 951
Db      301  TyrLysThrIleuArgAlaGluGlnIleAsnSerGlnIleValLysAsnThrMetThrGluThr 320
QY      952  CTGCTGTGTGACAGACCGCAACCCCGCATGTGCAAGACCATCTGCGCGCTTGGCGCGCGC 1011
Db      321  LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProGly 340
QY      1012  GCCACCTGGAGAGATATATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAAGGCC 1071
Db      341  AlaThrLeuGluGlnMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY      1072  CGCGTGTGTGCGCGAGCGCATGAGCGCGAG--GCCAACACCGTGAACATCATGATGAGAGA 1128
Db      361  ArgValIleuAlaGluAlaMetSerGlnIleThrAsnProAlaThrIleMetIleGlnLys 380
QY      1129  AGCAACTTCAAGAGCGCGCGCGCGCGCAAGTCAAGTCTTCAACTGCGCGCAAGAGAGGCCAC 1188
Db      381  GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnGlyLysGluGlnLys 400
QY      1189  ATGCGCAAGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248
Db      401  IleAlaLysAsnCysArgAlaProArgLysGlyCysTrpLysCysGlyLysGluGly 420
QY      1249  CACGATGATGAAGATGACCGAGCGCGCGCGCAACTCTGCGCGCAATCTGCGCGAGTGGCGCCAGC 1308
Db      421  HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY      1309  CACAAAGGCGCGCGCGCGCAACTTCTGCAAGAACCGCAGCGAGCGCGCGCGCGCGCGCGTG 1368
Db      441  HisLysGlyArgProGlyAsnAsnLeuGlnSerArgProGlu----- 454
QY      1369  CCCACCGCGCGCGCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCGCGC 1422
Db      455  ProThrAlaProProGluGlnIleSerPheArgPheGlyGluGlnThrThrThrProSerGln 474
QY      1423  AAGCAGAGCCCAAGAGCGCGAGCGCTTACCGCGAGCGCGCTGACCGCGCTGCGAGCGCTG 1482
Db      475  LysGlnGluProIleAspLysGluLeuTyr-----ProLeuAlaSerLeuArgSerLeu 492
QY      1483  TTGGCAGCGCGCGCGCTGAGCCAG 1506
Db      493  PheGlySerAspProSerSerGln 500

RESULT 5
US-09-718-096-16
; Sequence 16, Application US/09718096
; Patent No. 6589763
; GENERAL INFORMATION:
; APPLICANT: Von Jaer, Melke-Dorothee
; TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV
```

```

; FILE REFERENCE: 35-195
; CURRENT APPLICATION NUMBER: US/09/718,096
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: DE 19856463.5
; PRIOR FILING DATE: 1998-11-26
; PRIOR APPLICATION NUMBER: EP 99250415.9
; PRIOR FILING DATE: 1999-11-25
; PRIOR APPLICATION NUMBER: US 09/309,572
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: gag polyprotein
; US-09-718-096-16

```

## Alignment Scores:

Pred. No.:	1,42e-128	Length:	500
Score:	2246.00	Matches:	425
Percent Similarity:	90.16%	Conservative:	33
Best Local Similarity:	83.66%	Mismatches:	36
Query Match:	79.08%	Indels:	14
DB:	4	Gaps:	5

US-09-475-704a-4 (1-1509) x US-09-718-096-16 (1-500)

```

QY 1 ATGGGCGCGCGCGCGCATCTGCGCGCGAGAGAGCTGAGCAAGTGGAGAGATCCGC 60
DB 1 MetGlyAlaArgLaserValLeuSerGlyGlyLeuAspArgTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGAGAAAGCACTACATGCTGAACACCTGGTGGCCAGCCCGAG 120
DB 21 LeuArgProGlyGlyLysLysGlnTrpLysLeuLysIleValTrpAlaSerArgGlu 40
QY 121 CTGGAAGGCTTGGCCCTGGAACCCGCGCTGTGGAAGACCGCGAGGGCTGCAAGCATC 180
DB 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGlnThrSerGluLysCysArgGlnIle 60
QY 181 ATGAAGACCTGCAAGCCCGCCCTGCAAGACCGGCAAGAGAGCTGGCAAGCTTGAAC 240
DB 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuTrpAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGTCGACCGCGGCGATGAGGTCCGCAACACAGAGGCC 300
DB 81 ThrIleAlaValLeuTrpCysValIleGlnArgIleAspValLysAspThrLysGluAla 100
QY 301 CTGCAACAAGATCGAGAGAGAGAGACAAGTCCCAAGACAGAGACCCAGAGGCCAGAG 360
DB 101 LeuAspLysIleGluGlnGluGlnAsnLysSerLysLysValAlaGlnGlnAlaIle 120
QY 361 GCCGACGGC-----AAGGTAGCCAGAACTACCCCATCGTGTGAGAACTGCAAGGC 411
DB 121 AspThrGlyAsnAsnSerGlnValSerGlnAsnTrpProIleValGlnAsnLeuGlnGly 140
QY 412 CAGATGGTCAACCAAGCCATCAAGCCCGGCACTTGAAGCCCTGGGTGAAGGTGATCGAG 471
DB 141 GlnMetValIleGlnIleIleSerProArgThrLeuAsnAlaTrpValLysValValGlu 160
QY 472 GAGAAAGCTTGAAGCCCGGAGGTATCCCATGTTCAAGCCCTGAGCGAGGAGGCCAGCC 531
DB 161 GluLysAlaPheSerProGluValIleProMetPheSerIleAlaLeuSerGluGlyAlaTrp 180
QY 532 CCCAGGACCTGAACAAGATGTTGAACACCGTGGGCGGCCAGAGGCCAGCTGACAGATG 591
DB 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyIleGlnIleAlaIleMetGlnMet 200
QY 592 CTGAAGGACACCATCAACGAGAGGCGCGGAGTGGGACCGCTGACACCCCGTGCAGGCC 651
DB 201 LeuLysGluThrIleAsnGlnGluAlaIleGluTrpAspArgLeuIleAspValIleAla 220

```

```

QY 652 GCGCCCGTGGCGCGCGCGAGATGCGGACCCCGCGGAGAGACATGCGCGCGCAC 711
DB 221 GlyProIleAlaProGlyGlnMetArgGluProArgLysSerAspIleAlaGlyThrThr 240
QY 712 AGCACCTTGAGAGAGACAGATGCGCTGATGACCAAGAACCCCGCGTCCCGTGGCGGAC 771
DB 241 SerThrLeuGlnGlnGlnIleGlyTrpMetThrIleAsnProIleProValGlyGlu 260
QY 772 ATCTCAAGCGGTGGATGATCATCTGGCGCTGAAACAAAGATCGGGATGTACAGCCCGTG 831
DB 261 IleTrpLysArgTrpIleIleLeuGlyLeuAsnLysIleValIleArgMetTrpSerProThr 280
QY 832 AGCATCTGAGCATCGCGGAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTC 891
DB 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheAlaGlyPheValAspArgPhe 300
QY 892 TTCAAAGACCTTGGCGCGCGGAGCGGACCAAGAGAGCTGAAGAACTGGATGACCGAGACC 951
DB 301 TrpLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTCGAAGAACGCCCAACCCCGACTGCAAGAACATCTCGCGGCTTGGCGCGGCG 1011
DB 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProGly 340
QY 1012 GCCACCTTGAAGAGATGATGATACCGCTGCGGAGGCGTGGGCGCGCGCCAGCAAGGCC 1071
DB 341 AlaThrIleGluGlnMetMetThrIleCysGlnGlyValIleGlyProGlyIleIleValAla 360
QY 1072 CGCGTGTGGCGGAGCGCATGAGCCAG--GCCAAGACGTGAACATCATGATGACAGAG 1128
DB 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGlnLys 380
QY 1129 AGCACTTCAAGAGGCGCGCGCGCGCAACCTCAAGCTTCAACTCGCGGAAGAGGCGCAC 1188
DB 381 GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnGlyLysGlnGlyIle 400
QY 1189 ATGCGCAAGAACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCGGCAAGAGGAGC 1248
DB 401 IleAlaLysAsnGlyArgAlaProArgLysLysGlyStrLysCysGlyLysGlnLys 420
QY 1249 CACCAAGATGAAGACTGCAACGAGCGCGGCAAGCTTCTGGGCAAGATGCGCGCGCAC 1308
DB 421 HisGlnMetLysAspCysThrGluArgGlnIleAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAGAGGCGCGCGCGGCAACTCTCTGCAAGACGCAAGCGGACCGCGCGCGCCAGCGTG 1368
DB 441 HisLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
QY 1369 CCCACCGCGCGCGCGCGGAGAGCTTCCGCTTC-----GAGAGAGACCAACCCCGCGCCC 1422
DB 455 ProThrAlaProProGluGlnSerPheArgPheGlyGlnGluThrThrThrProSerGln 474
QY 1423 AAGCAGAGAGCCCAAGAGACCGGAGCTTACCGGAGCCCTGACCGCGCTTGGCAGCTTG 1482
DB 475 LysGlnGlnProIleAspLysGluLeuTrp-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTGCGCAGCGCGCGCGCTGAGCAG 1506
DB 493 PheGlySerAspProSerSerGln 500

```

## RESULT 6

US-09-952-060-35

Sequence 35 Application US/09952060

Patent No. 6733993

GENERAL INFORMATION:

APPLICANT: Emili, Emilio A.

APPLICANT: Youll, Rima

APPLICANT: Betz, Andrew J.

APPLICANT: Chen, Ling

APPLICANT: Kaslow, David C.

APPLICANT: Shiver, John W.

APPLICANT: Toner, Timothy J.

APPLICANT: Casimiro, Danilo R.

```

; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV-1 GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Codon optimized gag-1A pol fusion
; US-09-952-060-35

Alignment Scores:
Pred. No.: 5,01e-128 Length: 1350
Score: 2238.00 Matches: 423
Percent Similarity: 90.35% Conservative: 36
Best Local Similarity: 83.27% Mismatches: 35
Query Match: 78.80% Indels: 14
DB: 4 Gaps: 5

US-09-475-704a-4 (1-1509) x US-09-952-060-35 (1-1350)
QY 1 ATGGGCGCGCGCGCGCATCCGCGCGGCGGAGAGCTGAGACAGTGGAGAAATCCGC 60
DB 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGlnLeuAspTrpGlnLysIleArg 20
QY 61 CTGGCGCGCGCGCGGAGAGCACTACATGCTGAAGACACCTGATGGCGCGCGGAG 120
DB 21 LeuArgProGlyGlyLysValLysValLysValLysValLysValLysValLysVal 40
QY 121 CTGGAGGGCTTGCCCTGAACCCCGCGCTGCTGGAGACCGCGGAGGGCTGACAGCATC 180
DB 41 LeuGlnArgPheAlaValAsnProGlyLeuLeuGlnThrSerGlnGlyCysArgGlnIle 60
QY 181 ATGAAGACAGCTGACGCGCGCGCTGACGCGCGGACCGAGAGAGCTGGCGCTGTACAC 240
DB 61 LeuGlyGlnLeuGlnProSerLeuGlnTrpGlySerGlnGlnLeuArgSerLeuTrpAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGTGACGCGCGGATCGAGTCCGCGACCAAGAGGCC 300
DB 81 ThrValAlaThrLeuTrpCysValHisGlnLysIleAspValLysAspThrLysGlnAla 100
QY 301 CTGGACAAGATGAGAGAGAGAGCAAGCAAGTCCACAGAGAGACCAAGAGCGCAAGAG 360
DB 101 LeuGlnLysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
QY 361 GCGGACGGC-----AAGGTAGCGCAAGACTACCCGATGTGAGAACTGGAGCGGGC 411
DB 121 GlyThrGlyAsnSerSerGlnValSerGlnAsnTrpProIleValGlnAsnLeuGlnGly 140
QY 412 CAGATGCTCACCAGGCGCATCAGCCCGCGACCTGAGAGCGCTGGTGAAGTATCGAG 471
DB 141 GlnMetValHisGlnIleAlaIleSerProArgThrLeuAsnAlaIleTrpValLysValLysGln 160
QY 472 GAGAGAGCCTTACGCGCGGAGGTGATCCCATATTTCACGCGCTTGAGCGAGCGCCACC 531
DB 161 GlnLysAlaIlePheSerProGlnValIleProMetPheSerAlaLeuSerGlnGlyAlaThr 180
QY 532 CCCGAGAGCTGAGACGAGTGTGAACCGCTGGGGGGGCGACGAGCGCGCATGTGAGAG 591
DB 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnIleAlaIleMetGlnMet 200
```

```

QY 592 CTGAAGACACCATTAAGAGAGAGCGCGCGGAGTGGAGCCGCTGTACACCCCGTGCAGGCC 651
DB 201 LeuLysGlnThrIleAsnGlnGlnAlaIleGlnTrpAspArgLeuHisProValHisAla 220
QY 652 GGGCGCGGCGCGCGCGGAGTGGCGAGACCCCGCGGAGGAGATCGCGCGCGCGCAC 711
DB 221 GlyProIleAlaProGlyGlnMetCysGlnProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCTTGAGAGAGACAGATCGCTGATGACCAAGCAACCCCGCGTCCCGTGGGCGAC 771
DB 241 SerThrLeuGlnGlnGlnIleGlyTyrMetThrAsnAsnProProIleProValGlyGln 260
QY 772 ATCTCAAGCCGTGATCATCTGGGCTTGAACAAGATCGCGGATGTACAGCCCGGTG 831
DB 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProThr 280
QY 832 AGCATCTGAGACATCGCGCGGAGCGCGCGCAAGAGCGCTTCCGCGACTACGTGACCGCTTC 891
DB 281 SerIleLeuAspIleArgGlnGlyProLysGlnProPheArgPheTrpValAspArgPhe 300
QY 892 TTCAAGACCTTGCGCGCGCGAGAGGCGCAACGAGAGCTGAAGAACTGATGACCGAGAC 951
DB 301 TyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnTrpMetThrGlnThr 320
QY 952 CTGCTGTGTCAGAACCGCAACCCCGCATGTCAGACCATCTGCGCGCTTGCGCGCGCGC 1011
DB 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY 1012 GCCACCTTGAGAGAGATATGACCGCGCGGCGCGGCGCGCGCGCGCGCGCAAGGCC 1071
DB 341 AlaThrLeuGlnGlnMetMetThrAlaCysGlnGlyValGlyProGlyHisLysAla 360
QY 1072 CGCGTGTGCGCGGAGGCGATGAGCGAG--GCCAAGACGCTGAACATCATGTAGCAGAAG 1128
DB 361 ArgValIleAlaGlnAlaIleMetSerGlnValThrAsnSerAlaThrIleMetMetLysArg 380
QY 1129 AGCAACTTCAGAGGCGCGCGCGCGCAACGTCAGATGCTTCATCTGCGCGCAAGAGGCCAC 1188
DB 381 GlnAsnPheArgAsnGlnAlaGlyThrValLysCysPheAsnCysGlyLysValGlyHis 400
QY 1189 ATGGCGAAGACTGCGCGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCGGAGAGGGC 1248
DB 401 IleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGly 420
QY 1249 CACGAGATGAAGATGACCGAGCGCGCGCAACTTCTGGGCAAGATGTGCGCGCAC 1308
DB 421 HisGlnMetLysAspCysAsnGlnLysGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CAGAGGGCGCGCGCGCAACTTCTGAGAAACCGAGCGAGCGCGCGCGCGCACCGTG 1368
DB 441 HisLysGlyArgProGlyAsnPheLeuGlnSerArgProGln----- 454
QY 1369 CCCACCGCGCGCGCGCGGAGAGCTTCCGCTTC-----GAGGAGACCAACCCCGCGCCC 1422
DB 455 ProThrAlaProProGlnGlnSerPheArgPheArgGlyGlnGlyLysThrThrProSerGln 474
QY 1423 AAGCAGAGAGCCCAAGAGCGCGAGACCTACCGCGAGCCCTGTGACCGCGCTGGAGCGCTG 1482
DB 475 LysGlnGlnProIleAspLysGlnLeuTrp-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTTCGAGCGCGCGCGCTGAGCGGAC 1506
DB 493 PheGlyAsnAspProSerSerGln 500

RESULT 7
US-08-463-210-8
; Sequence 8, Application US/08463210
; Patent No. 6001977
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: WONG-STALL, Flossie
```







QY 1369 -----CCACCGCGCCCGCGAGAGCTTCCGC-----TTCCAG 1404  
 DB 461 LeuGlnSerArgProGluInuProThrAlaProProGluInuSerPheArgSerGlyAlaGlu 480  
 QY 1405 GAGACCAACCCCGCCCGCCCAAGAGAGCCGAGACCCCTACCGGAGCCCTG 1464  
 DB 481 ThrThrThrProGluInuSerGlnuProIleAerPlyGluLeuTyr-----ProLeu 498  
 QY 1465 ACCGCGCTGCGAGAGCTGTTCGAGCGCGCCCGCTGAGGCAG 1506  
 DB 499 ThrSerLeuArgSerLeuPheGlyAenAerProSerSerGln 512  
 RESULT 9  
 US-08-463-028-8  
 ; Sequence 8, Application US/08463028  
 ; Patent No. 6610476  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHANG, Nancy T.  
 ; APPLICANT: GALLO, Robert C.  
 ; APPLICANT: WONG-STAL, Flossie  
 ; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10154-0053  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,028  
 ; FILING DATE: 05-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 06/693,866  
 ; FILING DATE: 23-JAN-1985  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 06/659,339  
 ; FILING DATE: 10-OCT-1984  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Serutian, Leslie A.  
 ; REGISTRATION NUMBER: 35,353  
 ; REFERENCE/DOCKET NUMBER: 2026-4193US3  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 758-4800  
 ; TELEFAX: (212) 751-6849  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 512 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: HTLV-III  
 ; FEATURE:  
 ; NAME/KEY: Protein  
 ; LOCATION: 1..512  
 ; OTHER INFORMATION: /note="gag protein of HTLV-III"  
 ; US-08-463-028-8  
 Alignment Scores:  
 Pred. No.: 4.99e-128 Length: 512  
 Score: 2237.00 Matches: 422  
 Percent Similarity: 89.69% Conservative: 39  
 Best Local Similarity: 82.10% Mismatches: 39

Query Match: 78.77% Indels: 14  
 DB: 4 Gaps: 5  
 US-09-475-704a-4 (1-1509) x US-08-463-028-8 (1-512)  
 QY 1 ATGGGCGCCCGCGCGAGCATCTGCGCGGAGAGAGCTGACAAAGTGGAGAAATCCGC 60  
 DB 1 MetGlyAlaArgAlaSerValLeuSerGlyGluLeuAerArgTrpGluLeuArg 20  
 QY 61 CTGCGCCCGCGCGGAGAAAGCACTACATGCTGAAAGCACTGCTGTGGCCAGCCGCGAG 120  
 DB 21 LeuArgProGlyGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys 40  
 QY 121 CTGAGAGGCTTCGCGCCTGAAACCCCGCGCTGAGAGCCCGGAGGCTGCAAGCAGATC 180  
 DB 41 LeuGluArgPheAlaValAlaSerProGlyLeuLeuGluThrSerGlyGlyCysArgGlnIle 60  
 QY 181 ATGAAGCAGCTGAGACCCCGCGCTGAGAGCCGAGAGAGCTGCGAGCCTGTACAC 240  
 DB 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlyGlnLeuArgSerLeuTyrAsn 80  
 QY 241 ACCGTGGCCACCTGACTGCTGAGACCCGAGCATCGAGTCCGCGAGACCAAGAGGCC 300  
 DB 81 ThrValAlaThrLeuTyrCysValAlaSerGlnArgIleGluIleLysAspThrLysGluAla 100  
 QY 301 CTGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 DB 101 LeuAspLysIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 120  
 QY 361 GCCGAGCGC-----AAGGTGAGCCAGAACTACCCCATCTGTCAGAACCTGCGAGGC 411  
 DB 121 AspThrGlyLysLysSerSerGlnValSerGlnLeuThrProIleValGlnAlaMetGlnGly 140  
 QY 412 CAGATGTGTCACACAGAGCATCATGAGCCCGCGACCTGAAACGCTGTGGTGAAGTATCAG 471  
 DB 141 GlnMetValHisGlnAlaIleSerProArgThrLeuMetAlaTrpValLysValValGlu 160  
 QY 472 GAGAGGCGCTTCAGCGCCGAGAGTATCCCATGTTGACCGCGCTGAGAGAGAGCGCCAGC 531  
 DB 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180  
 QY 532 CCCGAGAGCTGAAACAGCATGTTGAACCGGTGGCGGAGCAAGCCGCGCATGAGATG 591  
 DB 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200  
 QY 592 CTGAAGGACACCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCC 651  
 DB 201 LeuLysGluThrIleAsnGluGluAlaIleGluTrpAspArgValHisProValHisAla 220  
 QY 652 GGGCGCGAGGCGCGCGCGAGATGGCGAGAGCCCGCGGAGAGCATGCGCGGCGCCAGC 711  
 DB 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240  
 QY 712 AGCACCCTGAGAGAGAGAGATGCGCTGATGACAGCAAGCCCGCGTGCCTGGGCGAGC 771  
 DB 241 SerThrLeuGlnGluGlnIleGlyTyrMetThrAsnAsnProProIleProValGlyLys 260  
 QY 772 ATCTCAAGCGGTGATATCTGTGGCTGAAACAAGATCGTGGGATTAACGCCCGCTG 831  
 DB 261 IleTyrLysArgTyrIleIleLeuGlyLeuAsnLysIleValAlaGlyMetTyrSerProThr 280  
 QY 832 AGCATCTGAGACATCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891  
 DB 281 SerIleLeuAspIleArgGlnGlyProLysGluLeuProPheArgAspTyrValAspArgPhe 300  
 QY 892 TTCAAGACCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCC 951  
 DB 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320  
 QY 952 CTGCTGTGTCAGAAAGCCCAAGCCGAGATGCAAGACCATCTCTCGGCGCTGCGCCGCGC 1011  
 DB 321 LeuLeuValGlnAsnAlaAsnProAspLysThrIleLeuLysValAlaLeuGlyProAla 340

OY	1012	GCACCCCTGGAGGACATATGACCCCTCCGACGAGCGTGGGCGGCCGACCAAGAGCC	1077
Db	341	AlaThrLeuGIuGlumetmetThrIaCysGlnGlyValGlyGlyProClyHisIySyaIa	360
OY	1072	CGCGTCTGGCCCGAGCGCATGAGCCAG--GCCAACAGCGTGAAATCATATGTCAGAAAG	1128
Db	361	ArgValLeuAlaGluAlaMetSerGlnValThrAenThrAlaThrIleMetetGlnArg	380
OY	1129	AGCAACTTCAAAGGCGCCCGGCGCAAGCGTCAAGGCTTCACATCGGCGCAAGAGGGCCAC	1188
Db	381	GlyAsnPheAlaGlnGlnArgGlySerValIySVCysPheAsnCYSGIyVlySgluIyHis	400
OY	1189	ATCGCGAAGAACTGCGCGCGCGCCCGCAGAAAGGCGCTGCTGAAGTGCGCAGAGAGGC	1248
Db	401	ThrIlaaArgAsnCysArgAlaProArgIySlySglYCySTrPlyCYSGIylySgluGly	420
OY	1249	CACCAAGATGAAGGACTGCAACGAGGCGCAGAGCCAACTTCGGGGAATCTGGGCGCACG	1308
Db	421	HisGlnmetLysASPcysThrGlnurGlnAlaAsnPheLeuGlyVlyIleTrpProSer	440
OY	1309	CACAAAGGCGCGCCCGGCGCAACTTCTCGAGAAACCGCAGCGAGCCCGCGCCCGCCACCGTG	1368
Db	441	TyrIySglYArgProGlyAsnPheLeuGlnSerArgProGluPuroThrAlaProPhe	460
OY	1369	-----CCACGCGCGCCCGCGCGCGAGAGCTTCGCG-----TTCCAG	1408
Db	461	LeuGlnSerArgProGluPuroThrAlaProProGluGlnSerPheArgSerGlyValGlu	480
OY	1405	GAGACCAACCCCGCGCCCGCAGAGCAGAGCCCAAGGACCGCGAGCCCTTACCGCAGCGCCGTG	1464
Db	481	ThrThrThrProProGlnIySgluGlnProIleAspIySgluLeuTyr-----ProIeu	498
OY	1465	ACCGCGCTGGCGAGCTGTTCGGGACGAGCGGCCCGCTCGAGCCAG	1506
Db	499	ThrSerLeuArgSerLeuPheGlyAsnAspProSerSerGln	512

RESULT 10  
 US-08-816-155B-45  
 Sequence 45, Application US/08816155B  
 Patent No. 5990091  
 GENERAL INFORMATION:  
 APPLICANT: TARTAGLIA, JAMES  
 APPLICANT: COX, WILLIAM I.  
 APPLICANT: GETTIG, RUSSELL R.  
 APPLICANT: MARTINEZ, HECTOR  
 APPLICANT: PAOLETTI, ENZO  
 APPLICANT: PINCUS, STEVEN E.  
 TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND  
 TITLE OF INVENTION: METHODS OF MAKING AND USING THEREOF  
 NUMBER OF SEQUENCES: 48  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FROMMER LAWRENCE & HAUG LLP  
 STREET: 745 FIFTH AVENUE  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10151  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/816,155B  
 FILING DATE: 12-MAR-1997  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KOMALSKI, THOMAS J.  
 REGISTRATION NUMBER: 32,147  
 REFERENCE/DOCKET NUMBER: 454310-2990  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-588-0800  
 TELEFAX: 212-588-0500

```

; INFORMATION FOR SEQ ID NO: 45
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-816-155B-45

```

**Alignment Scores:**

Pred. No.:	3.51e-127	Length:	500
Score:	2223.00	Matches:	420
Percent Similarity:	90.16%	Conservative:	38
Best Local Similarity:	82.68%	Mismatches:	36
Query Match:	78.27%	Indels:	14
DB:	2	Gaps:	5

US-09-475-704A-4 (1-1509) x US-08-816-155B-45 (1-500)

[illegible]

```

Db      261  ILeTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetYrSerProthr 280
QY      832  AGCATCTGAGACATCCGCGCAGAGGCGCCCAAGAGACCTTCGCGACATCGAGCGGCTTC 891
Db      281  SetIleuAsnPrIleArgInGlyProLysGlyProPheArgAspTyrValAspArgPhe 300
QY      892  TTCAGACCCCTGGCGCGCGAGACGCGCACCGACAGAGCTGAGAACTGATGACCGAGCC 951
Db      301  TyrLysThrLeuArgLysIleGlnIleAsnSerGlnIleValLysAsnTrpMetThrGln 320
QY      952  CTGCTGTGTGACGAGACCGCAACCCCGGACTGCAACACCATCTGCGCGCTCTCGCGCCG 1011
Db      321  LeuLeuValGlnIleAsnLysAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 340
QY      1012  GCCACCTCTGAGAGAGATGATGACCGCTGCGACGAGGCGGTGGCGCGCGCGCGCAAG 1071
Db      341  AlaThrLeuGlnIleMetMetThrAlaCysGlnGlyValGlyProGlnIleLysAla 360
QY      1072  CCGGTGCTGGCGCGCGAGATGACCGAG---GCCACAGCGGTGAACATCATGATGACAG 1128
Db      361  ArgValIleuAlaGlnIleMetSerGlnIleValThrAsnSerAlaThrIleMetMet 380
QY      1129  AGCAATTCAGAGGCGCGCGCGCGCAACGTCAAGTCTTCACTGCGCGCAAGAGGCGCAC 1188
Db      381  GlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnGlyGlyLysGlnGly 400
QY      1189  ATGCGCAAGAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGAAAGTGGCGCAAGAGGCG 1248
Db      401  ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGln 420
QY      1249  CACCGAGTAAAGATGTCACCGAGCGCGCAAGCGCAACTTCCTGGGCAAGATCTGGCGC 1308
Db      421  HisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpPro 440
QY      1309  CACAAGGGCGCGCGCGCGCAACTCTGCGCAAGACCGGAGCGCGCGCGCGCGCGCGCG 1368
Db      441  TyrLysGlyArgProGlnIleAsnPheLeuGlnSerArgProGln----- 454
QY      1369  CCCACCGCGCGCGCGCGCGAGAGCTTCGCG-----TTGAGAGAGACACCGCGCGCG 1422
Db      455  ProThrAlaProProGlnIleLysPheArgSerGlyValGlnThrThrProProGln 474
QY      1423  AAGCAGAGAGCCCAAGAGACCGGAGCCCTTACCGGAGCGCGCGCGCGCGCGCGCGCG 1482
Db      475  LysGlnGlnProIleAspLysGlnLeuTyr-----ProLeuThrSerLeuArgSerLeu 492
QY      1483  TTCGGAGGGCGCGCGCGCGCGAG 1506
Db      493  PheGlyAsnAspProSerSerGln 500

RESULT 11
US-09-079-587-45
; Sequence 45. Application US/09079587
; Patent No. 6130066
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PROLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079, 587
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/816,155
; FILING DATE: 12-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454310-2990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-09-079-587-45

Alignment Scores:
Pred. No.: 3,51e-127 Length: 500
Score: 2223.00 Matches: 420
Percent Similarity: 90.16% Conservatve: 38
Best Local Similarity: 82.68% Mismatches: 14
Query Match: 78.27% Indels: 16
DB: 3 Gaps: 5

US-09-475-704A-4 (1-1509) x US-09-079-587-45 (1-500)
QY      1  ATGGCGCGCGCGCGCGAGCATCTGCGCGCGAGAGCTGAGCAAGTGGAGAGATCCGC 60
Db      1  MetGlyAlaArgAlaSerValLeuSerGlyGlnLeuAspArgTrpGlnLysIleArg 20
QY      61  CTGGCGCGCGCGCGCGCAAGAGCATGCTGAAAGCACTGTGTGGCGCGCGCGAG 120
Db      21  LeuArgProGlyGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys 40
QY      121  CTGGAGGGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db      41  LeuGlnArgPheAlaValAsnProGlyLeuLeuGlnTrpSerGlnGlyCysArgGln 60
QY      181  ATGAAGCAGCTGACCGCGCGCGCGCGCGAGACCGGACCGAGAGCTGCGAGCTGTAC 240
Db      61  LeuGlnIleuGlnProSerLeuGlnThrGlySerGlnIleuLysSerLeuTyrAsn 80
QY      241  ACCGTGGCACCCTGTACTGCTGTCGACCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
Db      81  ThrValAlaTrpLeuTyrCysValHisGlnArgIleGlnIleLysAspThrLysGln 100
QY      301  CTGGCAAGATCGAGAGAGAGAGAGAGCAAGTCCCGCAGAGAGACCGCAGAGGCAAG 360
Db      101  LeuAspLysIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 120
QY      361  GCCAGCGC-----AAGGTAGCCAGAACTACCCATCGTGCAGAACCTGCGAGGCG 411
Db      121  AspThrGlnLysSerAsnGlnValSerGlnAsnTrpProIleValGlnAsnIleGln 140
QY      412  CAGATGTGACACCAAGCCATACGCGCGCACCTGAAAGCTGAGTGTGAGGTATCGAG 471
Db      141  GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValGln 160
QY      472  GAGAGGGCTTCAGCGCGCGAGATGATCCCATGTTTACCGCGCTGAGAGGCGCGCAC 531
Db      161  GlnLysAlaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnGlyAla 180
QY      532  CCCAGGACCTGAACACAGATGTGAAACCGTGGCGCGCGCACGAGCGCGCATGAGATG 591

```

```

Db      181  ProGlnAspLeuAenThrMetLeuAenThrValGlyHisGlnAlaIleMetGlnMet 200
Qy      592  CTGAAGAGAACCATCAACAGAGAGAGCCCGGAGTGGAGCCGCTGACCCCGTGAAGGCC 651
Db      201  LeuLysGluThrThrIleAsnGlnGlnAlaIleGlnTTPAspArgValHisProValHisAla 220
Qy      652  GGCCTGCTGGCCCGGAGAGATGCGCAGACCCCGGCGAGAGACATCCCGGCGCAGC 711
Db      221  GlyProIleAlaProGlyGlnMetArgGlnProArgGlySerAspIleAlaGlyThrThr 240
Qy      712  AGCACCCTGAGAGAGAGATGCGCTGATGACCAAGACCCCGCTGCGCTGGCGCAG 771
Db      241  SerThrLeuGlnGlnGlnIleGlyTrpMetThrAsnAsnProProIleProValGlyGlu 260
Qy      772  ATCTACAAACGGGTGATCATCTGGGCTGGAACAGATGTGGATGATGACCCCGG 831
Db      261  IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
Qy      832  AGCATCTGATCATCGCAGAGGCGCCCAAGAGCCCTTCGCGAGCTGAGACGCTTC 891
Db      281  SerIleLeuAspIleArgGlnGlyProLysGlnProAspArgPyrValAspArgPhe 300
Qy      892  TTCAGAGACCTGCGCGCCGAGAGAGCCACCAAGAGCTGAGAAATGATGACCGAGACC 951
Db      301  TyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnTrpMetThrGlnThr 320
Qy      952  CTGCTGCTGCAAGAGCCCAACCCGAGCTGGAAGACCATCTGCGCTGCGCCCGCGC 1011
Db      321  LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
Qy      1012  GGCACCCCTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGGCGCCCAAGAGCC 1071
Db      341  AlaThrLeuGlnGlnMetThrAlaCysGlnGlyValGlyProIleHisLysAla 360
Qy      1072  CCGCTGCTGGCCGAGGCGAGTGAAGCCAG--GCCAAGAGCGTGAACATCATGATGAGAG 1128
Db      361  ArgValIleuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetGlnArg 380
Qy      1129  AGCACTTCAAGAGGCGCCCGGCGCAAGTCAAGTGTCACTGCGGCGCAAGAGGCGCAC 1188
Db      381  GlnAsnPheArgAsnGlnAlaArgLysIleValLysCysPheAsnGlyGlyLysGlnGlyHis 400
Qy      1189  ATCCGCAAGAAATGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGGCAAGAGGCC 1248
Db      401  ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGly 420
Qy      1249  CACCAAGATGAAGATGACACGAGCGCAGGCGCAACTTCCTGGGCAAGATGCGCCAGC 1308
Db      421  HisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
Qy      1309  CACAAGGCGCGCCCGGCAACTTCTGCAAGAACCGGAGGAGCCCGCCCGCCACCGTG 1368
Db      441  TyrLysGlyArgProGlyLysPheLeuGlnSerArgProGlu----- 454
Qy      1369  CCCACGCGCGCCCGCGGAGAGCTTCGCG-----TTGAGAGAGACCAACCCCGCGCCC 1422
Db      455  ProThrAlaProProGlnGlnLeuSerPheArgSerGlyValGlnThrThrProProGln 474
Qy      1423  AAGCAGAGAGCCCAAGAGCGGAGCCCTACCGGAGCCCTGAGCGCGCTGGGAGCGCTG 1482
Db      475  LysGlnGlnProIleAspLysGlnLeuTyr-----ProLeuThrSerLeuArgSerLeu 492
Qy      1483  TTCGCGAGCGCGCCCTGAGCAG 1506
Db      493  PheGlyAsnAspProSerSerGln 500

```

```

/ APPLICANT: Youil, Rima
/ APPLICANT: Bet, Andrew J.
/ APPLICANT: Chen, Ling
/ APPLICANT: Kaslow, David C.
/ APPLICANT: Shiver, John W.
/ APPLICANT: Toner, Timothy J.
/ APPLICANT: Casimiro, Danilo R.
/ TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
/ TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
/ TITLE OF INVENTION: MODIFICATIONS
/ FILE REFERENCE: 20747Y
/ CURRENT APPLICATION NUMBER: US/09/952,060
/ CURRENT FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: PCT/US01/28861
/ PRIOR FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: 60/317,814
/ PRIOR FILING DATE: 2001-09-07
/ PRIOR APPLICATION NUMBER: 60/279,056
/ PRIOR FILING DATE: 2001-03-27
/ PRIOR APPLICATION NUMBER: 60/233,180
/ PRIOR FILING DATE: 2000-09-15
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 33
/ LENGTH: 493
/ TYPR: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: cpa-gag fusion open reading frame
/ US-09-952-060-33

Alignment Scores:
Pred. No.: 2,29e-119 Length: 493
Score: 2094.00 Matches: 396
Percent Similarity: 90.53% Conservative: 34
Best Local Similarity: 83.37% Mismatches: 31
Query Match: 73.73% Indels: 14
DB: 4 Gaps: 5

US-09-475-704a-4 (1-1509) x US-09-952-060-33 (1-493)
Qy      100  CTGATGTGGAGCGAGCGGAGTGGAGGCTTCGCCCTGAACCCCGGCTGCTGAGAGCC 159
Db      27  IleValThrAlaSerArgGlnLeuGlnArgPheAlaValAsnProGlyLeuLeuGlnThr 46
Qy      160  GCCGAGGCTGCAAGACATGATGAGAGCGTGCAGCCGCGCTGCAAGCGGACCGCAG 219
Db      47  SerGlnGlyCysArgGlnIleLeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlu 66
Qy      220  GAGCTGCGGAGCTGTACAAACCGTGGCCACCTGTACTGCGTGCAGCCCGCATGAG 279
Db      67  GlnLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnLysIleAsp 86
Qy      280  GTCCGCGACACCAAGAGAGCCCTGGAAGATGAGAGAGGAGCGAAGAACTCCAGCAG 339
Db      87  ValLysAspThrLysGlnAlaLeuGlnLysIleGlnGlnGlnAsnLysSerLys 106
Qy      340  AAGACCCAGGCGCCCAAGAGAGCGCGC-----AAGTGAAGCCAGAACTACCCC 390
Db      107  LysAlaGlnGlnAlaAlaIleGlyThrGlyLysAsnSerSerGlnValSerGlnAsnTrpPro 126
Qy      391  ATCTGCAAGAACTTGACAGGCGCAAGTGTGACCAAGCCATCAGCCCGCAGCCTTGAC 450
Db      127  IleValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsn 146
Qy      451  GCTGCGTGAAGTGAAGTGAAGAGAGCCCTTCAGCCCGCGAGTGAATCCCATGTTCACC 510
Db      147  AlaTrpValLysValValGlnGlnLysValAlaPheSerProGlnValIleProMetPheSer 166
Qy      511  GCCCTGAGCGAGGCGCCACCCCGCAGGACCTGGAACAGATGTGAACACCGTGGCGCGC 570
Db      167  AlaLeuSerGlnGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGly 186

```

RESULT 12  
 US-09-952-060-33  
 ; Sequence 33, Application US/09952060  
 ; Patent No. 6733993  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eminti, Emilio A.

QY	571	CACGACGGCCGCATGAGATGCTGAAAGACAACATTAAGAGAGGCCGCCGAGTGGAGC	630
Db	187	HISGlnAlaIaIaMeClnMeTLeuYsGlnTrrIleAsnGlnGlnAlaIaIaGlnTrrPaaP	206
QY	631	CCCTGACACCCCGTCGACGGCCGGCCCGTCGGGCCCGCCGACATGGGACACCCCGCGAC	690
Db	207	ArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMeCArgGlnProArgGly	226
QY	691	AGCGACATGTCGGCGGCCACCAAGCACCTTCGACAGAGACAGATGCGCTGATGACCAACAC	750
Db	227	SerAspIleAlaGlyThrThrSerThrLeuGlnGlnGlnIleGlyYrMetThrAsnAsn	246
QY	751	CCCCCGCGTCCCGTCGGGCGACATCTAACAGCGGTGATCATCTGGGCGCTGAACAAGATC	810
Db	247	ProProIleProValGlyIuIleYrLysArgTrrPileIleLeuGlyLeuAsnYsIle	266
QY	811	GTGCGGATGTACAGCCCGCTGAGACATCTTGAGCATCGCGACAGGGCCCAAGAGAGCCCTTC	870
Db	267	ValArgMetTyrSerProThrSerIleLeuAspIleArgGlnGlyProYsGlnProPhe	286
QY	871	CGCGACTACGTGACCGCTTCTTCAAGACCTTCGCGCGACAGAGGCCACCAAGACGTCG	930
Db	287	ArgAspYrYrValAspArgPheYrYrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnVal	306
QY	931	AAGAATCGATATACCGAGACCCCTGCTGCTGGTCAGAACGCCAACCCCGCATCTGCAAGACATC	990
Db	307	LysAsnTrrMetThrGlnThrIleLeuValGlnAsnAlaAsnProAspCysLysThrIle	326
QY	991	CTGCGCGCTCTGCGCCCGCGCGCACCTTCGTGAGAGATGATGACCGCTTCGACAGGCGCTG	1051
Db	327	LeuYrValIaLeuGlyProAlaAlaThrIleLeuGlnGlnMeCethrAlaCysGlnGlyAla	346
QY	1051	GCGCGCCCGCGGCACAAAGGCCCGCGCTGTCGCGACGACATGAGCCAGC--GCCAACAC	1107
Db	347	GlyGlyProGlyYrIleYrValIaArgValIleuAlaGlnAlaIaMeSerGlnValThrAsnSer	366
QY	1108	GTGAACATCATATGTCGAAAGAGCAACTTCAAGAGGCCCGCGCGCAAGCTCAAGTCTTC	1167
Db	367	AlaThrIleMeMetGlnArgGlyAsnPheArgAsnGlnArgYrThrValIysCysPhe	386
QY	1168	AACTCGCGCAAGAGGGCCACATCGCCCAAGACTGCGCGGCCCGCGCAAGAAAGGGCTGC	1227
Db	387	AsnCysGlyLysValGlyHisIleAlaLysAsnCysArgAlaProArgYrLysGlyCys	406
QY	1228	TGGAGTGGCGCAGAGGGGCCACCAAGATGAAGAGATGTGACCGAGCGCCAGGCCCACTTC	1287
Db	407	TrpYrCysGlyLysGlyGlnGlyIleGlnIleMetLysAspCysAsnGlnIaArgIleAsnPro	426
QY	1288	CTGGGCACAGATTTGGCCCAAGCCACCAAGAGGGCGCGCCCGCACTTCTTCGAAACCGCAGC	1347
Db	427	LeuGlyLysIleTrrProSerHisYsGlyArgProGlyYrAsnProLeuGlnSerArgPro	446
QY	1348	GAGCCCGCGGCCCAACCGTGCACACCGCGCCCGCCCGCGCAAGACTTCGCGTTC-----	1407
Db	447	Glu-----ProThrAlaProProGlnGlnIleSerThrPheGlyGln	460
QY	1402	GAGGAGACCAACCCCGCCCGCCCAAGCAGAGACCAAGACGACCGCAAGCCCTTACCGCAGCC	1467
Db	461	GlyYrThrThrProSerGlnLysGlnGlnProIleAspLysGlnIleuYr-----Pro	478
QY	1462	CTGACCGCGCTCGCGACGCTTCTTCGCGACGGCGCCCTGAGCCAG	1506
Db	479	LeuAlaSerLeuArgSerLeuPheGlyAsnAspProSerSerGln	493

```

NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,499A
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/102/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEX: (202)672-5399
FAX: 904136
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-08-127-499A-11

Alignment Scores:
Pred. No.:      8.57e-115      Length:      478
Score:          2018.50       Matches:     382
Percent Similarity: 87.86%    Conservative: 45
Best Local Similarity: 78.60%   Mismatches:  40
Query Match:      71.07%      Indels:     19
DB:               1           Gaps:        5

US-09-475-704A-4 (1-1509) x US-08-127-499A-11 (1-478)
QY      1 ATGGGCGCCCGCGGCAGCATCTGTGCGCGGCGAGAAGCTGTGAACAAGTGGAGAGATTCGC 60
Db      1 MetGlYAlArghLaSerValLeuSerGIyglIuLeuApArGrTpGlUyRlEArg 20
QY      61 CTGCGCGCCGCGGCGAAGACACTCAATGCTGAAGACCTGTGTGTGGCCAGCCGCCGAG 120
Db      21 LeuArPrGdLyglYleYleYleYrTyRlyLeuYshLileVAtTPALASerArgLu 40
QY      121 CTGAAGAGGCTTGCCCTGTAAACCCCGGCTGTGAGAACCGCGAGGCGTGAAGACATC 180
Db      41 LeuGIuArghPheAlVaIAsnProdlYeuLeuGIuThrsErGIuGLyCySRgGlnIlle 60
QY      181 ATGAAGCAGGTGACGCCGCGCCCTGTGAAGACCGGCAACGAGAGCTGSCAGCTGTACAC 240
Db      61 LeuGIyGIuLeuGIuNproSerLeuGIuThrgISerGIuGIuLeuArGrSerLeuYrAsn 80
QY      241 ACCGTGGCAACCTGTACTGTCTGTGACGCCCGGCAATCGAGTCCGCGACCAAGAGGCC 300
Db      81 ThvAlAlArthrIeuTrICyEvAlHtSgInArGIleGIuILeuySAepThrLYSGluAla 100
QY      301 CTGAACAAGATCGAGGAGGAGCAACAAGTCCGAGAGAGACCAGACCCAGAGCCAAAGAG 360
Db      101 LeuApPuylIleGIuSIuGIuInmenYserLySvlYbSaIagInGlnAlaAla 120
QY      361 GCCAGCGC-----AAGGTGACCAAGAACTAACCCCATCTGTGAACAACCTGACAGGC 411
Db      121 AspThrGIyHisSerSerclnValSerGIuInenTYrProIIeValGlnAnIllegInCIy 140
QY      412 CAGATGGTGACCAAGGCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTATGAG 471
Db      141 GlmErValHtSgInAlalIesErProArGrThrIeuAmnlATPrVAllyVAlvaIGlu 160

```

QY 472 GAGAGGCGCTTCAGCCCGGAGGATCCCATGTCACCGCCCTGAGCGGCGGCGCACC 531  
 DB 161 GILYVALAIPheserProgluValIlePrometPheserIleuSerGluValIleThr 180  
 QY 532 CCCAGACCTGAAACAAGATGTTGAACACCGTGGGCGGCACAGCGCCGATGACATG 591  
 DB 181 ProGlnApeuAhnThrMetIleuAhnThrValGlyGlyHisGlnIleAhnMetGlnMet 200  
 QY 592 CTGAAGAACACCATCAACAGAGAGCGCGCGAGTGGACCGCTGACACCCCGTGACGCC 651  
 DB 201 LeuLyGluThrIleAhnGluGlnIleAhnGluIleAhnGluIleAhnGluIleAhnGluIle 220  
 QY 652 GGGCCCGTGGCCCGCGGAGATCCCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 711  
 DB 221 GlyProIleAhnProIleAhnProIleAhnProIleAhnProIleAhnProIleAhnProIle 240  
 QY 712 AGCAGCTGAGAGAGATCCGCTGATGACAGCAACCCCGCGTGGCGGCGGCGGCGGCGG 771  
 DB 241 SerThrLeuGlnGluIleGlyIleThrMetThrAhnAhnProIleProValGlyGln 260  
 QY 772 ATCTCAAGCGGTGATCATCTGCGGCGTGAACAGATCGTGGAGTGAACAGCCCGT 831  
 DB 261 IleTyIleAhnGlyIleIleIleGluIleAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIle 280  
 QY 832 AGCATCTGAGACATCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 891  
 DB 281 SerIleAhnAhnProIleAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIle 300  
 QY 892 TTCAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 951  
 DB 301 TyrIleThrLeuAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIle 320  
 QY 952 CTGCTGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1011  
 DB 321 LeuLeuValGlnAhnAhnAhnProAhnProAhnProAhnProAhnProAhnProAhnProAhn 340  
 QY 1012 GCCACCTGAGAGAGATGATGACCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1071  
 DB 341 AlaThrLeuGlnGluIleAhnMetThrIleAhnGlyIleAhnGlyIleAhnGlyIleAhnGlyIle 360  
 QY 1072 CGCGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1128  
 DB 361 ArgValLeuAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 380  
 QY 1129 AGCAATTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1188  
 DB 381 GlyAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 400  
 QY 1189 ATGCGCCAAAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1248  
 DB 401 ThrAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 420  
 QY 1249 CACCAAGATGAAGATGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1305  
 DB 421 HisGlnMetLeuAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 440  
 QY 1306 AGCCAAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1365  
 DB 441 ThrArgGluGlyIleGlnIleIlePheAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 455  
 QY 1366 GTGCCCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1425  
 DB 456 -----GlnProHisIlePheAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 466  
 QY 1426 CAGAGCCCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1443  
 DB 467 GlnGlnProHisGlnIle 472  
 RESULT 14  
 US-08-482-847-11  
 ; Sequence 11, Application US/08482847  
 ; Patent No. 5556757

GENERAL INFORMATION:  
 APPLICANT: VAN ALSTYNE, Diane  
 APPLICANT: SHARMA, Lawrence Rajendra  
 TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
 BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
 TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
 NUMBER OF SEQUENCES: 40  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/482,847  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/127,499  
 FILING DATE: 28-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 51916/104/INBI  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 672-5300  
 TELEFAX: (202) 672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 478 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 US-08-482-847-11  
 Alignment Scores:  
 Pred. No.: 8.57e-115 Length: 478  
 Score: 2018.50 Matches: 382  
 Percent Similarity: 87.86% Conservative: 45  
 Best Local Similarity: 78.60% Mismatches: 40  
 Query Match: 71.07% Indels: 19  
 DB: 1 Gaps: 5  
 US-09-475-704A-4 (1-1509) x US-08-482-847-11 (1-478)  
 QY 1 ATGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60  
 DB 1 MetGlyAlaArgAlaSerValIleuSerGlyGluLeuAhnAhnAhnAhnAhnAhnAhnAhnAhn 20  
 QY 61 CTGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
 DB 21 LeuAhnProGlyGlyIleAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 40  
 QY 121 CTGAGAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
 DB 41 LeuGluAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 60  
 QY 181 ATGAAGAGCTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
 DB 61 LeuGlyGlnLeuGlnProSerIleuGlnThrIleAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 80  
 QY 241 ACCGTGCGCACCTGTAAGTGTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 DB 81 ThrValAlaThrLeuTyIleValHisGlnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhnAhn 100  
 QY 301 CTGCAAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

```

Db      |||||||
101 LeuAapLyIleGlUGlUGlInAenLySerLyLyLeValaInGlnIlaIaIaIa 120
QY      361 GCCGAGCGC-----AAGTGAGCCGAAGACTACCCATGCTGCAAGAACCTGCAGCGGC 411
Db      121 AapThGlyHIsaSerGInValSerGInAenLyProIleValGlnAenIleGInGly 140
QY      412 CAGATGGTCAACAGGCGCATCAGCCCGCCGCAACCTCAAGCGCGGTGGAAGGATCGAG 471
Db      141 GlnMetValHIsGlnIlaIleSerProAgnThLeuAenHlaIaTrpValIyValaIaGln 160
QY      472 GAGAAAGCTTCAAGCCCGCAGGTGATCCCATGTTTCAACCGCCCTGAGCAGGCGCCAC 531
Db      161 GlnLyValaIaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnGlyAlaThr 180
QY      532 CCCCAGGACCTGAACACGATGTTGAACACCGTGGGGCGCCACAGGCCCGCCATGCAATG 591
Db      181 ProGlnAapLeuAenHlaenThMetLeuAenThrValGlyGlyHIsGlnIlaIaMetGlnMet 200
QY      592 CTGAAGGACCATCAACGAGGAGCGCGCAGTGGAGCGCGCTGCACCCCGTCAAGCGC 651
Db      201 LeuLySerGlnThIleAenGlnGlnIlaIaGlnIlaIaGlnIlaIaIaIaIaIa 220
QY      652 GGCCTCGTGGCCCGCCGCAAGATGCGCGACCCCGCGCGCAGCAGCATCGCCGCGCGCAC 711
Db      221 GlyProIleAlaProGlyGlnMetAaGlnProAaGlySerAaPylleAaGlyThrThr 240
QY      712 AGCACCCTGCAAGAGAGATCGCTGATGACAGACACCCCGCTGCGCGCGCGCAC 771
Db      241 SerThleuGlnGlnIleGlnIleGlyTrpMetThraAaAenProIleProValaGlyGln 260
QY      772 ATCTAAGAGCGGTGATCATCTCTGGGCTGAACAAGATGATGATGATGATGATGATGAT 831
Db      261 IleTrLyLyAaGTrIleIleIleuGlyLeuAaenLyIleValaAaGlySerProThr 280
QY      832 AGCATCTGGAACATCGCGCAGCGCGCCCAAGAGCCCTTCGCGCATCTGACCGCTTC 891
Db      281 SerIleLeuAaPylleAaGlnGlyProLyGlnProPheAaPylleAaPylleAaPylle 300
QY      892 TTCAGAGACCTCGCGCCGCGCAGCGCGCCAGCGCGCATGTAAGAACCTGATGACCGAG 951
Db      301 TyrLySerThleuAaGlnIleGlnIleAaSerGlnIleValaAaAenThrMetThrGln 320
QY      952 CTGCTGCTGCAAGACCCCAACCCCGCTCAAGACCATCTGCGCGCTCTGCGCGCGCGC 1011
Db      321 LeuLeuValaGlnAaHlaAaenProAaPylleThrIleuLyAaIaLeuGlyProAla 340
QY      1012 GCCACCTTGAGAGATGATGACCGCTGCGCGCGCGCGCGCGCGCGCGCAAGAGCC 1071
Db      341 AlaThrLeuGlnGlnIleuMetThraIaCySgInGlyValaGlyProGlyHIsLyVala 360
QY      1072 CCGTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1128
Db      361 ArgValaLeuAaIaGlnIleuMetSerGlnValaThraAaThraIaThrIleuMetSerGln 380
QY      1129 AGCAACTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1188
Db      381 GlyAaenPheAaGlnIleuAaGlnIleuAaGlnIleuAaGlnIleuAaGlnIleuAa 400
QY      1189 ATCGCCAGAACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248
Db      401 ThrAlaAaGlnAaenLyAaGlnIleuAaGlnIleuAaGlnIleuAaGlnIleuAa 420
QY      1249 CACCGATGAAGATCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1305
Db      421 HIsGlnMetLyAaPylleThrGlnIleuAaAaenPheLeuGlyLyIleCyAaLeuPro 440
QY      1306 AGCACAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1365
Db      441 ThrAaGlnGlnIleuGlnIleuPhePheAaGlnIleuAaGlnIleuAaGlnIleuAa 455
QY      1366 GTGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1425

```

```

Db      456 -----GlnProHIsIlePhePheAaGlnIleuAaGlnIleuAaGlnIleuAa 466
QY      1426 CAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1443
Db      467 GlnGlnProHIsGlnIleuAaGlnIleuAaGlnIleuAaGlnIleuAaGlnIleuAa 472

RESULT 15
US-08-392-794A-2
; Sequence 2, Application US/08392794A
; Patent No. 6025141
; GENERAL INFORMATION:
; APPLICANT: HU, Yu-Wen
; TITLE OF INVENTION: IMMUNOFLUORESCENCE ASSAY FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES USING RECOMBINANT ANTIGENS IN INSOLUBLE FORM
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,794A
; FILING DATE: 09-DEC-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,789
; FILING DATE: 10-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: NEIMARK, SHERIDAN
; REGISTRATION NUMBER: 20,520
; REFERENCE/DOCKET NUMBER: HU=4A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-392-794A-2

Alignment Scores:
Pred. No.: 2,41e-114 Length: 437
Score: 2011.00 Matches: 374
Percent Similarity: 93.14% Conservative: 33
Best Local Similarity: 85.58% Mismatches: 26
Query Match: 70.81% Indels: 4
DB: 3 Gaps: 2

US-09-475-704A-4 (1-1509) x US-08-392-794A-2 (1-437)
QY      1 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db      1 MetGlyAlaAaGlnAaSerValLeuSerGlyGlyGlnLeuAaPylleAaGlnIleuAa 20
QY      61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db      21 LeuAaGlnProGlyGlyLyLyLyLyLyLyLyLyLyLyLyLyLyLyLyLyLyLyLyLy 40
QY      121 CTGAGGCGCTTCCCGCTGAACCGCGCGCTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCG 180
Db      41 LeuGlnAaGlnPheAaIaAaenProGlyLeuLeuGlnIleuAaGlnIleuAaGlnIleuAa 60
QY      181 ATGAAGAGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

```



Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGlnLeuArgSerLeuTyrAsn 80  
QY 241 ACCGTGGCCACCTGTACTGCTGTGCACGGCCGATGAGGTCCGCACCAAGAGGCC 300  
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgGlnGlnLeuIleLysAspThrIysGlnAla 100  
QY 301 CTGGAACAAGATGAGAGGAGGAGAACAGTCCACAGACCAAGAACCCACAGAGCCCAAGAG 360  
Db 101 LeuAspLysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120  
QY 361 GCCGACGGC-----AAGGTGAGCCAGAACTACCCCATGTCGAGAACCTGCAGAGCC 411  
Db 121 AspThrGlnHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140  
QY 412 CAGATGTCGACACGAGCCATCAGCCCGCCGACCTGTGACGCTGAGGTGAAGTATGAG 471  
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTyrValLysValValGln 160  
QY 472 GAGAAGGCTTGAAGCCCGAGGTGATCCCATGTTACACCGCCCTGAGCGAGGGCCGAC 531  
Db 161 GlnLysAlaPheSerProGlnValIleProMetPheSerAlaLeuSerGlnGlnAlaThr 180  
QY 532 CCCAGAGACCTGAACACGATGTTGAACACGTTGGGCGGCCACCAAGGCCGACATGACATG 591  
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMet 200  
QY 592 CTGAAGACACCATCAACAGAGAGCGCGCGAGTGGACCGCTGCACCCCGTGCAGGCC 651  
Db 201 LeuLysGlnThrIleAsnGlnGlnAlaIleGlnTyrAspArgValHisProValHisAla 220  
QY 652 GGCCCCGTGGCCCCCGGCGAGATGGCGGACCCCGGCGAGGACATCCCGCGGCCAC 711  
Db 221 GlyProIleAlaProGlyGlnMetArgGlnProArgGlySerAspIleAlaGlyThrThr 240  
QY 712 AGCACCTGACAGAGAGATGCGCTGATGACACGACACCCCGCGTGGCGGCGAC 771  
Db 241 SerThrLeuGlnGlnGlnIleGlyTyrMetThrAsnAsnProIleProValGlyGln 260  
QY 772 ATCTACAGCGGTGATCATCTGTGGGCTGAACAGATGTGCGGATGTACAGCCCGCTG 831  
Db 261 IleTyrLysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280  
QY 832 AGCATCTGTGACATCCGCCAGAGGCCGCCAAGAGCCCTTCCGGACTACGTGACCGCTTC 891  
Db 281 SerIleLeuAspIleArgGlnGlyProLysGlnProPheArgAspTyrValAspArgPhe 300  
QY 892 TTCAGAGACCTGCGCGCCGAGCAGGCCACCGACGATGAAGACTGATGACCGAGACC 951  
Db 301 TyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnThrMetThrGlnThr 320  
QY 952 CTGCTGTGTCAGAACGCGCAACCCCGCATGCAAGACCATCTGCGCGCTTCGCGCCCGGC 1011  
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340  
QY 1012 GCCACCTGTGAGAGATGATGACCGCTTCCAGGGGTGGGGCGGCCGCCCAAGGCC 1071  
Db 341 AlaThrLeuGlnGlnMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360  
QY 1072 CCGGTGCTGGCGAGCGATGAGCCAG--GCCAAGCGGTGAACATCATGATGACAGAG 1128  
Db 361 ArgValLeuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380  
QY 1129 AGCAACTTCAAGAGGCCCGCGGCAACGTCAAGTGTTCACCTGCGGCAAGAGAGGCCAC 1188  
Db 381 GlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGlnGlnHis 400  
QY 1189 ATCGCCAAAGAACTGCGCGGCCGCCGCAAGAGAGGCTGCTGGAAGTGGCGCAAGAGGCC 1248  
Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGlnGly 420  
QY 1249 CACCAATGAAGATGACGACCGAGCGCCAGGCCCAACTCTGCGGCAAGATC 1299

Db 421 HisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlnLysIle 437

Search completed: September 1, 2005, 23:19:54  
Job time : 48.261 secs

**This Page Blank (usp10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_nzp model

Run on: September 1, 2005, 22:09:18 ; Search time 386.088 Seconds  
(without alignments)  
3078.466 Million cell updates/sec

Title: US-09-475-704A-4

Perfect score: 2840

Sequence: 1 atggggcggccgcgcagcatc.....ggcggccctgagccagtraa 1509

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 3548624

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+npz.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US09475704/rnatc\_01092005\_160752\_19548/app\_query.fasta\_1.3342  
-DB=Published Applications AA -OPT=fastran -SUFFIX=xapb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USRP=US09475704\_@CGN\_1\_1\_233\_@rnatc\_01092005\_160752\_19548  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5  
-FGAPOF=6 -FGAPEXT=7 -YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.dep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.dep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.dep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.dep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.dep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.dep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.dep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.dep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.dep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.dep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.dep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09D\_NEW\_PUB.dep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.dep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.dep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.dep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.dep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.dep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.dep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11B\_PUBCOMB.dep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11C\_NEW\_PUB.dep:\*  
21: /cgn2\_6/ptodata/1/pubpaa/US11D\_NEW\_PUB.dep:\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.dep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2661	93.7	502	10	US-09-899-575-22
2	2467	86.9	492	18	US-10-780-507-97
3	2464	86.8	492	16	US-10-332-413-2
4	2464	86.8	492	18	US-10-494-131-36
5	2457	86.5	492	9	US-09-991-258-5
6	2457	86.5	492	18	US-10-929-234-5
7	2457	86.5	493	16	US-10-332-413-8
8	2456.5	86.5	493	18	US-10-780-507-98
9	2450	86.3	492	18	US-10-780-507-96
10	2440	86.9	492	14	US-10-339-217-109
11	2430	85.6	492	14	US-10-339-217-110
12	2421.5	85.3	491	14	US-10-339-217-140
13	2413.5	85.0	491	14	US-10-339-217-143
14	2387	84.0	508	14	US-10-339-217-105
15	2385.5	84.0	491	14	US-10-339-217-104
16	2375.5	83.6	497	14	US-10-339-217-107
17	2353.5	82.9	487	14	US-10-339-217-106
18	2348	82.7	496	14	US-10-339-217-108
19	2343	82.5	492	10	US-09-899-575-17
20	2263	79.7	500	18	US-10-780-507-47
21	2258.5	79.5	1228	18	US-10-494-131-30
22	2258	79.5	500	14	US-10-339-217-18
23	2251	79.3	442	18	US-10-494-131-32
24	2247.5	79.1	494	14	US-10-339-217-126
25	2246	79.1	500	14	US-10-223-172A-42
26	2246	79.1	500	14	US-10-224-999A-3483
27	2245	79.1	500	18	US-10-780-507-48
28	2240.5	78.9	497	16	US-10-800-023-21
29	2238	78.8	1350	10	US-09-952-060-35
30	2238	78.8	1350	16	US-10-380-641-35
31	2238	78.8	1350	17	US-10-636-730-35
32	2237	78.7	512	17	US-10-634-165-8
33	2234	78.7	500	18	US-10-780-507-49
34	2232	78.6	557	14	US-10-339-217-144
35	2226	78.4	494	16	US-10-335-168-8
36	2225.5	78.4	494	14	US-10-339-217-100
37	2223	78.3	500	14	US-10-097-534-32
38	2223	78.3	500	14	US-10-059-271-85
39	2223	78.3	500	14	US-10-102-622-4
40	2223	78.3	500	16	US-10-325-168-34
41	2223	78.3	500	17	US-10-844-658-6
42	2221	78.2	500	9	US-09-968-355-17
43	2221	78.2	583	9	US-09-968-355-17
44	2220.5	78.2	499	17	US-10-844-658-2
45	2220	78.2	498	16	US-10-325-468-22

## ALIGNMENTS

RESULT 1  
US-09-899-575-22  
; Sequence 22, Application US/09899575  
; Publication No. US20030223961A1  
; GENERAL INFORMATION:  
; APPLICANT: Zur Megele, Jan  
; APPLICANT: Barnett, Susan W.  
; APPLICANT: Egnelbrecht, Susan  
; APPLICANT: van Rensburg, Estrelita Janse  
; TITLE OF INVENTION: POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USBS THEROF  
; FILE REFERENCE: PP01631.102  
; CURRENT APPLICATION NUMBER: US/09/899,575  
; CURRENT FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 09/475,704  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 502

```

/ TYPE: PRF
/ ORGANISM: Human immunodeficiency virus
US-09-899-575-22

Alignment Scores:
Pred. No. : 9,17e-129      Length: 502
Score: 2661.00      Matches: 502
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 93.70%      Indels: 0
DB: 10      Gaps: 0

US-09-475-704a-4 (1-1509) x US-09-899-575-22 (1-502)

QY 1 ATGGGGCGCGCGCGCGCGCATCTCGCGCGCGGAGAACTGTGACAACTGGGAGAAAGATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyGluValLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACCTGGTGTGGCCAGCCGCGAG 120
DB 21 LeuArgProGlyGlyLysLysHisTrpMetLeuLysHisLeuValITrpAlaSerArgGlu 40
QY 121 CTGAGAGGCTTGCCTGTGAACCCCGGCTGTGTGAGACCGCGGAGGGCTGTCAAGCATC 180
DB 41 LeuGluGlyPheAlaLeuAsnProGlyLeuLeuGluThrAlaGluGlyCysLysGlnIle 60
QY 181 ATGAAGACACTGACGGCGCGCGCGCGCATCGGACCGGACGAGAGCTGGCGAGCTGTACAC 240
DB 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluLeuArgSerLeuLysAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGTGACGCGCGGCGCATGAGGTCGCGGACCAAGAGAGCC 300
DB 81 ThrValAlaThrLeuLysCysValHisAlaGlyIleGluValAlaArgAspThrLysGlnAla 100
QY 301 CTGGACAAGATCGAGAGAGAGCAAGAACAGTCCACGACAGAGAACCCAGAGGCCAAAGAG 360
DB 101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnIleAlaLysGln 120
QY 361 GCGGACGGGAGGTGAGCGAAGACTACCCCATGTGTGACAGCTGTGAGGCGGACAGATGTC 420
DB 121 AlaAspGlyLysValSerGlnAsnLysTrpIleValGlnAsnLeuGlnLysMetVal 140
QY 421 CACGAGCCCATGAGCCCGCGCACCTGTGAAGCTGTGGGTGAAGGTGATCGAGAGAGAGCC 480
DB 141 HisGlnAlaIleSerProArgThrLeuAsnAlaITrpValLysValIleGluGluLysVala 160
QY 481 TTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGGCGAGGGCGCCACCCCGGAGAC 540
DB 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACAGATGTTGAACACCGTGGGGGGCGGACGAGCGCGCATGTGAGATGCTGAAGAGAC 600
DB 181 LeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGAGACCGGCTGACACCCGTCGAGGCGGCGCCGTC 660
DB 201 ThrIleAsnGluGlnAlaAlaGluTrpAspArgLysHisProValGlnIleAlaGlyProVal 220
QY 661 GCCCGCGCGCGAGTGGCGGACCCCGCGGAGGAGCATGCGCGGCGCCACGAGACCTTG 720
DB 221 AlaProGlyGlnMetArgAspProArgLysSerAspIleAlaGlyAlaThrSerTrpLeu 240
QY 721 CAGGACAGATGCGCTGGATGACCAAGAACCCCGCGTCCGCTGGGGCGACATCTTCAAG 780
DB 241 GlnGluGlnIleLeuAlaTrpMetTrsAsnProProValProValLysAspIleLys 260
QY 781 CGGTGATCATCTGGGCGCTGAAACAAGATCGTGGGATGACAGCCCGCTGAGCATCTG 840
DB 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProValSerIleLeu 280
QY 841 GACATCGCGCAGGGCGCCAAAGAGAGCCCTTCCGAGTACGTGAGCCGCTTTTCAAGACC 900
DB 281 AspIleArgGlnGlyProLysGlnProPheArgAspTyrValAspArgPhePheLysThr 300
```

```

QY 901 CTGCGCGCGGAGGAGGCGCACCCAGGACGTGAAGACTGTGATGACGAGACCTGTGCTG 960
DB 301 LeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrGluThrLeuLeuVal 320
QY 961 CAGAACGCGCAACCCCGCATGTGCAAGACATCTGTGCGGCTCTCGGCGCGCGCGCACCTG 1020
DB 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGGATGATGACCGCTGCCAGGAGCGTGGCGGCGCGCGCGCACAGGCGCGCTGCTG 1080
DB 341 GlnGluMetThrAlaCysGlnGlyValAlaLysProGlnHisLysAlaArgValLeu 360
QY 1081 GCGGAGGAGATGAGCGGACGCGCAACGAGCTGAACATCAAGATGACGAGAACGAACTTCAAG 1140
DB 361 AlaGluAlaMetSerGlnAlaAsnSerValAsnIleMetMetGlnLysSerAsnPheLys 380
QY 1141 GCGCGCGCGCGCAACGTCAAGTCTTCAACTGTGCGCAAGAGGCGCACATCGCCAAAGAC 1200
DB 381 GlyProArgArgAsnValLysCysPheAsnCysGlyLysGluHisIleAlaLysAsn 400
QY 1201 TGCGCGCGCGCGCGCAAGAGGCTGTGAAAGTGTGCGCAAGAGGCGCACCAATGAAG 1260
DB 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlnHisGlnMetLys 420
QY 1261 GACTGCACCGGAGCGCGCGGCGCAACTCTGTGGCAAGATCTGCGCGCACGACCAAGGCGCGC 1320
DB 421 AspCysTrpGlnArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGCGCAACTTCTGTGACAAACGCGAGCGGCGCGCGCGCGCACCGTGCACGCGCGCC 1380
DB 441 ProGluAsnPheLeuGlnAlaAsnArgSerGluProAlaAlaProThrValProThrAlaPro 460
QY 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGACGACCCCGCGCGCGCGCGCGCGCGCGCGCGC 1440
DB 461 ProAlaGlnSerPheArgPheGluGluThrThrProAlaProLysGlnGluProLysAsp 480
QY 1441 GCGGAGCCCTACCGGAGACCCCTGACCGCGCTGTGCGACACTGTTCGCGAGGCGCGCGCTG 1500
DB 481 ArgGluProTrpArgGluProLeuThrAlaLeuArgSerLeuPheGlySerGlyProLeu 500
QY 1501 AGCCAG 1506
DB 501 SerGln 502

RESULT 2
US-10-780-507-97
; Sequence 97, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MILLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780, 507
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 492
; TYPE: PRF
```

```

; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Least squares center of tree reconstruction of clade C gag protei
; US-10-780-507-97
Alignment Scores:
Pred. No.: 8, 41e-119 Length: 492
Score: 2467.00 Matches: 466
Percent Similarity: 95.62% Conservative: 12
Best Local Similarity: 93.23% Mismatches: 12
Query Match: 86.87% Indels: 10
DB: 18 Gaps: 2

US-09-475-704a-4 (1-1509) x US-10-780-507-97 (1-492)

QY 1 ATGGGCGCCGCGCGAGCATCTGCGCGCGAGAGAGCTGAGCAAGTGGAGAGATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyLysLeuAspThrTrpGluSerIleArg 20
QY 61 CTGCGCGCCGCGCGAGAGAGCACTGATGCTGAGACCTGTGTGGCCAGCCCGGAG 120
DB 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisLeuValTrpAlaSerArgGlu 40
QY 121 CTGAGAGGCGCTGCGCCCTGAACCCCGCGCTGTGAGAGACCGCGAGGGGCTGCAAGCATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlySerGlnIle 60
QY 181 ATGAGCAGCTGAGAGCCCGCGCTGAGAGACCGCGAGAGCTGTGCGAGCTGTCAAC 240
DB 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTCTGATCTGCTGCGACGCCGGGATGAGGTCCGCGACCAAGAGGCC 300
DB 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100
QY 301 CTGAGCAAGATCGAGAGAGAGAGCAAGAGTCCAGCAGAGAGCCAGAGGCGAGAG 360
DB 101 LeuAspLysIleGluGlnGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaGluAla 120
QY 361 GCCGAGCGAGAGGTGAGAGAGAGAGAGTCCCATCTGTCAGAGAGCTGAGAGGCGCATGTG 420
DB 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMetVal 140
QY 421 CACCGAGCATGAGAGCCCGCGAGCCTGAAAGCTGTGGTGAAGTGAAGTGAAGAGCC 480
DB 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
QY 481 TTCAGCGCGAGAGTATCCCATGTTCACCGCGCTGAGAGAGAGGCGCGAGCCCGCGAGAG 540
DB 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTTGAACACCGTGGGCGGCGACACGAGCGCGCATGAGATGCTGAAGAG 600
DB 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACAGAGAGCGCGCGAGTGGAGACCGCTCACCCCGCTGACAGCGCGCGCGT 660
DB 201 ThrIleAsnGluGluAlaIleGluTrpAspArgLysHisProValHisAlaGlyProVal 220
QY 661 GCCCGCGCGAGATGGCGAGCCCGCGCGAGAGATAGCGCGCGCGCGCGAGAGCCG 720
DB 221 AlaProGlyGlnMetArgLysProAlaGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGCAGATCGCTGTGATGACACAGAACCCCGCGTGGCGCGGAGCATCTTCAAG 780
DB 241 GlnGluGlnHisAlaIleTrpMetThrSerAsnProProValProValGlySerIleTyrLys 260
QY 781 CGGTGATCATCTCTGGGCTGAAACAAGATGTGTGGATGTACAGCCCGTGAAGATCTTG 840
DB 261 ArgTrpIleIleLeuGlyLeuLeuLysIleValArgMetLysSerProValSerIleLeu 280
QY 841 GACATCGCGAGGCGCGCGAGAGAGCCCTTCCGAGACTAGTGAGCCGCTTTCAGAGCC 900

```

```

DB 281 AspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPheLeuSerThr 300
QY 901 CTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 301 LeuArgAlaGluGlnAlaIleThrGlnAspValLysAsnTrpMetThrAspThrLeuVal 320
QY 961 CAGAACGCCAACCCCGAGCTGCAAGAGACATCTCTGCGCGCTTGGGCGCGCGCGAGAG 1020
DB 321 GlnAsnAlaAsnProAspLysSerThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGAGATGATGACCGCGCTGCGAGGCGCGTGGGCGCGCGCGAGAGAGAGAGAGAGAG 1080
DB 341 GlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLeuValArgValLeu 360
QY 1081 GCCGAGCGATGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 361 AlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgSerAsnPheLys 380
QY 1141 GCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 381 GlyProLysArgIleValLysCysPheAsnLysGlyLysGluGlyHisIleAlaArgAsn 400
QY 1201 TCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 401 CysArgAlaProArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
QY 1261 GACTCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 AspCysThrGluArgGlnAlaAsnAsnLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGCGCACTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 441 ProGluAsnPheLeuGlnSerArgProGlu-----ProThrLapro 454
QY 1381 CCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 455 ProAlaGluSerPheArgPheGluGlnThrThrProAlaProLysGlnLysProLysAsp 474
QY 1441 CGCGAGCGCTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
DB 491 SerGln 492

RESULT 3
US-10-332-413-2
; Sequence 2, Application US/10332413
; Publication No. US20040116660A1
; GENERAL INFORMATION:
; APPLICANT: Johnstone, Robert Edward
; APPLICANT: Swanson, Ronald Ivar
; APPLICANT: Morris, Lynn
; APPLICANT: Karim, Salim Abdool
; APPLICANT: Williamson, Carolyn
; TITLE OF INVENTION: Process for the Selection of HIV-1 Subtype C Isolates, Selected n
; FILE REFERENCE: 45669-281993
; CURRENT APPLICATION NUMBER: US/10/332,413
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: PCT/IB01/01208
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,995
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: ZA 2000/3437
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: ZA 2000/4924
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2

```

```
/ LENGTH: 492
/ TYPE: PRF
/ ORGANISM: Human immunodeficiency virus type 1
US-10-332-413-2

Alignment Scores:
Pred. No.: 1,2e-118 Length: 492
Score: 2464.00 Matches: 466
Percent Similarity: 95.62% Conservative: 14
Best Local Similarity: 92.83% Mismatches: 12
Query Match: 86.76% Indels: 10
DB: 16 Gaps: 2

US-09-475-704a-4 (1-1509) x US-10-332-413-2 (1-492)

QY 1 ATGGGCGCGCGCGCGCATCTCGCGCGCGGAGAGCTGGAACAAGTGGGAGAAATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGGCAAGAGCACTACATGCTGAAGCACTGGTGTGGCCAGCGCGAG 120
DB 21 LeuArgProGlyGlyLysLysHisTrpMetLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGGAAGGCTTGCGCCCTGAACCCCGCGCTGTGAGAGACCGCGAGGCTTGCAAGCATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATGAAGAGAGCTGACGCGCGCGCTGACAGCGCGGACCGAGAGAGCTGGCGAGCTTACAC 240
DB 61 MetLysGlnLeuGlnProAlaLeuGlnTrpGlyThrGluLeuLysSerLeuTyraen 80
QY 241 ACCGTGGCCACCTGACTGCTGCTGACGCGCGCATGAGGCTCGGACACCAAGAGAGCC 300
DB 81 ThrValAlaThrLeuTrpCysValHisGlyLysIleGluValArgAspThrLysGlnAla 100
QY 301 CTGGAACAAGATGAGAGAGAGAGAGAAACAAGTCCACAGAGAGAGAGAGAGAGAGAG 360
DB 101 LeuAspLysIleGluGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCGGACGCGGAGAGTGAAGAGAACTACCGCATGCGAGAGAGAGAGAGAGAGAGAGAG 420
DB 121 AlaAspGlyLysValSerGlnAsnTrpProIleValGlnAsnLeuGlnIleMetVal 140
QY 421 CACAGAGCCATCAGCGCGCGCACCCCTGAAGCGCTGGGTGAAGTATCGAGAGAGAGCC 480
DB 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIleGluGluLysAla 160
QY 481 TTCAGCGCGGAGGTGATCCCATGTTCAACCGCGCTGAGAGAGAGAGAGAGAGAGAG 540
DB 161 PheSerProGluValIleProMetCpHeThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACAGAGATGTTGAACAACGCGGCGCGGCGCACAGGCGCGCATGAGATGCTGAAGAG 600
DB 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGAGACCGCTGACACCGCGTGAAGCGCGCGCGG 660
DB 201 ThrIleAsnGluGluAlaIleGluTrpAspArgValHisProValHisIleGlyProIle 220
QY 661 GCGCGCGCGGAGATGGCGAGACCCCGCGGAGCGAGATGCGCGGCGCGCACAGACCTTG 720
DB 221 AlaProGlyGlnMetArgLysTrpArgLysSerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGGAGAGATGCGCTGGATGACCAAGCAACCCCGCGTGCCTGGGGGCACTTACAAAG 780
DB 241 GlnGluGlnIleAlaIleTrpMetThrSerAsnProProIleProValGlyAspIleTrpLys 260
QY 781 CGGTGGATCATCTGGGCTTGAACAAGATGTCGGATATACAGCCCGCTGAGATCTCGG 840
DB 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProValSerIleLeu 280
QY 841 GACATCCGCGAGAGGCGCGCAAGAGAGCGCTTCCGCGACTACGTGAGAGCGCTTCTCAAGAC 900
```

```
DB 281 AspIleArgGlnGlyProLysGluProPheArgAspTrpValAspArgPhePheLysThr 300
QY 901 CTGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 301 LeuArgAlaGluGlnAlaThrGlnIleValLysAsnTrpMetThrAspThrLeuVal 320
QY 961 CAGAACGCGCAACCCCGAGTGAAGACATCTGCGCGCTCGAGCCCGGCGGAGAGAGAG 1020
DB 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGGAGATGATGACCGCTGCGGAGGAGCGTGGCGGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 341 GlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgValLeu 360
QY 1081 GCGGAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 361 AlaGluAlaMetSerGlnAlaAsnSerGlyAsnIleMetMetGlnArgSerAsnLys 380
QY 1141 GCGCGCGCGCGCAACGTCAGTCTTCAACTGCGGCAAGAGAGAGAGAGAGAGAGAGAG 1200
DB 381 GlyProArgArgIleValLysCysPheAsnGlyGlyLysGluGlyHisIleAlaArgAsn 400
QY 1201 TGCGCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisIleGlnMetLys 420
QY 1261 GACTGACCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGGCAACTTCTGCAAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454
QY 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGACACCCCGCGCGCGAGAGAGAGAGAGAGAGAGAG 1440
DB 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProIleGlu 474
QY 1441 GCGGAGCGCTACCGGAGAGCGCGCTGACGCGCGTGGGAGAGAGAGAGAGAGAGAGAG 1500
DB 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
DB 491 SerGln 492

RESULT 4
US-10-494-131-36
/ Sequence 36, Application US/10494131
/ Publication No. US20050176929A1
/ GENERAL INFORMATION:
/ APPLICANT: Williamson, Carolyn
/ APPLICANT: Van Harmelen, Joanne Heidi
/ APPLICANT: Gray, Clive Maurice
/ APPLICANT: Bourn, William
/ APPLICANT: Karim, Salim Abdool
/ TITLE OF INVENTION: HIV-1 Subtype Isolate Regulatory/Accessory Genes, and
/ TITLE OF INVENTION: Modifications and Derivatives Thereof
/ FILE REFERENCE: 45669-300571
/ CURRENT APPLICATION NUMBER: US/10/494,131
/ CURRENT FILING DATE: 2004-04-30
/ PRIOR APPLICATION NUMBER: PCT/IB02/04550
/ PRIOR FILING DATE: 2002-10-31
/ PRIOR APPLICATION NUMBER: ZA 2001/8978
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 36
/ LENGTH: 492
/ TYPE: PRF
/ ORGANISM: Human immunodeficiency virus type 1
US-10-494-131-36
```

## Alignment Scores:

Pred. No.:	1,2e-118	Length:	492
Score:	2464.00	Matches:	466
Percent Similarity:	95.62%	Conservative:	14
Best Local Similarity:	92.83%	Mismatches:	12
Query Match:	86.76%	Indels:	10
DB:	18	Gaps:	2

US-09-475-704a-4 (1-1509) x US-10-494-131-36 (1-492)

```

QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGGAGAGCTGGACAAAGTGGAGAAATCCGC 60
DB 1 MetG1A1Aarg1Aser11e1eua1rg1G1u1Lys1eua1s1p1r1p1u1s11e1a1rg 20
QY 61 CTGGCGCCCGCGCGCAAGGCACTACATCTGAGCAGCTGGTGGGCGCCAGCCGCGAG 120
DB 21 Leua1r1p1ro1y1g1Lys1u1s1A1s1T1r1Met1eul1h1s1l1e1a11r1p1A1s1e1a1rg1u 40
QY 121 CTGGAGGGCTTCCGCGTGAACCCCGGCGTGTGGAGACCGCGGAGGGGCTGGCAAGATC 180
DB 41 LeuG1u1A1r1p1h1A1e1u1a1n1p1ro1y1Lys1eul1u1h1s1e1r1u1g1Lys1s1g1n1l1e 60
QY 181 ATGAACAGCTGAGCGCCGCGCTGACAGCCGCGACCGAGAGGCTGGCGACCTGTACAAC 240
DB 61 Met1y1G1u1eul1n1p1ro1A1eul1n1h1r1G1Y1h1r1G1u1eul1y1s1e1r1eul1y1r1a1n 80
QY 241 ACCGTGGCCACCTGTACTGCTGTCAGCGCGGATGAGGTCCGCGACACCAAGAGGCC 300
DB 81 Thr1Val1A1a1h1r1eul1y1r1Cys1Val1h1s1G1u1y1s1l1e1G1u1A1A1r1g1A1s1p1h1r1y1s1G1u1A1 100
QY 301 CTGGACAAGTGGAGAGGAGGAGCAAGTCCACAGAGGCCAGAGCGCCAGAGAG 360
DB 101 Leua1s1p1y1s1l1e1G1u1G1u1n1a1n1Lys1Cys1G1n1Lys1Thr1G1n1A1A1y1s1A1 120
QY 361 GCGGAGCGCAAGGTGAGCGCAAGTACCCCATGTCAGAACCTTGAGGCGCCAGATGTG 420
DB 121 A1a1a1s1p1G1y1Lys1e1s1e1r1n1eul1y1r1p1ro1L1e1Val1G1n1h1n1eul1n1G1y1Met1eul1 140
QY 421 CACGAGGCCATCAGCCCGCGACCTGAAAGCCTGGGTGAGGTATCGAGAGAGGCC 480
DB 141 H1s1G1n1A1A1l1e1s1e1r1p1ro1A1r1g1Th1r1Leua1n1A1A1r1p1Val1y1s1A1l1e1G1u1n1y1s1A1 160
QY 481 TTCAGCGCCGAGGTATCCCAATGTTACCGCGCTGAGAGGAGGCGCCACCCCGCGAGC 540
DB 161 P1h1e1s1e1r1p1ro1G1u1A1l1e1p1ro1Met1p1h1r1A1A1e1s1e1r1u1y1A1A1h1r1p1ro1G1n1A1p 180
QY 541 CTGAACACGATGTTGAACAACCGTGGCGCGCACAGCGCGCCATGCGAGATGCTGAAGAC 600
DB 181 Leua1n1Th1r1Met1eul1n1h1r1Val1G1y1G1h1s1G1n1A1A1A1Met1G1n1Met1eul1y1s1A1p 200
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGAGACCGCTGACCCGCGTGCAGGCGCGCGCTG 660
DB 201 Thr1L1e1A1n1G1u1n1A1A1A1G1u1T1r1P1A1s1p1A1r1g1Val1h1s1p1ro1A1h1s1A1A1G1y1P1ro1L1e 220
QY 661 GCGCGCGCGCAATGGCGACCCCGCGCGCGAGATAGCGCGCGCCACACAGACCTG 720
DB 221 A1A1p1r1o1G1y1G1n1eul1y1r1g1u1p1ro1A1r1g1y1s1e1r1A1s1l1e1A1G1y1h1r1h1s1e1r1Th1r1Leu 240
QY 721 CAGGACCAATCGCTGATGACCAACAACCCCGCGTGGCGCGCGCGCGCAATCTTCAAG 780
DB 241 G1n1G1u1n1h1l1e1A1A1r1p1Met1h1r1s1e1r1A1n1p1ro1L1e1p1ro1A1G1y1s1p1l1e1T1r1y1s 260
QY 781 CCGTGATCATCTGGGCTTGAACAAGATGTGCGGATGTACAGCCCGCTGAGACCTCTG 840
DB 261 Arg1T1r1p1L1e1l1eul1y1Lys1eul1y1s1l1e1Val1A1r1G1Met1y1s1e1r1p1ro1A1s1e1r1l1e1u 280
QY 841 GACATCGGAGGCGCGCAAGAGGCGCTTCCGCGATTAAGTGAACCGCTTCTTCAAGAC 900
DB 281 Asp1L1e1A1r1g1n1L1y1p1ro1y1s1u1p1ro1p1h1e1A1r1g1A1s1p1h1r1p1h1eul1y1s1T1r 300
QY 901 CTGCGCGCGAGAGGCGCAACCGAGAGCTGAAGAACTGATGACCAAGACCTGTGCGTG 960
DB 301 Leua1r1g1A1G1u1n1A1A1h1r1G1n1G1u1Val1y1s1A1n1T1r1Met1h1r1A1s1p1h1r1eul1eul1 320

```

```

QY 961 CAGAAAGCCAAACCCGAGCTGCAAGACATCTTGGCGCTTGGCGCGCGCGCCACCTCG 1020
DB 321 G1n1A1n1A1A1s1p1ro1A1s1p1y1s1Th1r1L1e1u1A1A1eul1y1p1ro1y1A1A1h1r1Leu 340
QY 1021 GAGGATGATGACCGCGCTGCGAGGCGCGGCGCGCGCGCGCGCGCAAGGCGCGCGTG 1080
DB 341 G1n1G1u1n1eul1Met1h1r1A1A1Cys1G1n1G1y1A1G1y1p1ro1G1y1h1s1y1A1A1r1g1Val1eul1 360
QY 1081 GCGGAGCGGATGAGCGAGCGCAACAGCGTGAACATCATGATGACAGAGCAATTTCAAG 1140
DB 361 A1A1G1u1A1A1e1s1e1r1G1n1A1A1n1s1e1r1G1y1A1n1L1e1u1eul1Met1G1n1A1r1s1e1n1h1eul1y1s 380
QY 1141 GCGCGCGCGCGCAAGTCAAGTCTTCAACTGCGCGCAAGAGGCGCGCAATCCCAAGAAC 1200
DB 381 G1y1p1ro1A1r1g1L1e1Val1y1s1Cys1p1h1e1n1Cys1G1y1s1G1u1n1h1s1l1e1A1A1A1r1g1A1n 400
QY 1201 TGCAGGCGCGCGCGCAAGAGGCGTGTGAAAGTGTGGCAAGAGGCGCGACAGATGAAG 1260
DB 401 Cys1A1r1g1A1A1p1ro1A1r1y1s1G1y1s1y1r1p1y1s1G1y1y1s1G1u1n1h1s1l1e1Met1y1s 420
QY 1261 GACTGACCGAGCGCGCGCAAGCTTCTTGGCGCAAGATCTGGCGCGACAGCGCGCGCG 1320
DB 421 Asp1Cys1Th1r1G1u1A1r1G1n1A1A1n1h1eul1y1s1l1e1T1r1p1ro1s1e1r1h1e1y1s1G1y1A1r1g 440
QY 1321 CCGCGCACTTCTTGAAGACCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 441 P1ro1y1A1n1p1h1eul1G1n1A1n1A1r1p1ro1G1u1-----P1ro1Th1r1A1p1ro 454
QY 1381 CCGCGCGAGAGCTTCCGCTTGGAGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB 455 P1ro1A1A1G1u1s1e1r1p1h1eul1G1u1h1r1Th1r1p1ro1A1p1ro1y1s1G1u1n1p1ro1L1e1G1u 474
QY 1441 CCGGAGCCCTTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
DB 475 -----Arg1u1p1ro1eul1h1s1e1r1eul1y1s1e1r1eul1h1p1h1eul1y1s1e1r1p1ro1eul1 490
QY 1501 AGCCAG 1506
DB 491 Ser1G1n 492

```

RESULT 5  
US-09-991-258-5  
Sequence 5, Application US/09991258  
Patent No. US20020141975A1  
GENERAL INFORMATION:  
APPLICANT: Olmsted, Robert  
APPLICANT: Keith, Paula  
APPLICANT: Dryga, Sergey  
APPLICANT: Caley, Ian  
APPLICANT: Maughan, Maureen  
APPLICANT: Johnston, Robert  
APPLICANT: Davis, Nancy  
APPLICANT: Swanson, Ronald  
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE  
FILE REFERENCE: 01113.0001U3  
CURRENT APPLICATION NUMBER: US/09/991,258  
PRIORITY FILING DATE: 2001-11-16  
PRIORITY FILING DATE: 2001-07-09  
PRIORITY FILING DATE: 2001-07-09  
PRIORITY FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 492  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =  
US-09-991-258-5

## Alignment Scores:

Pred. No.:	2,74e-118	Length:	492
Score:	2457.00	Matches:	465
Percent Similarity:	95.22%	Conservative:	13
Best Local Similarity:	92.63%	Mismatch:	14
Query Match:	86.51%	Indels:	10
DB:	9	Gaps:	2

US-09-475-704a-4 (1-1509) x US-09-991-258-5 (1-492)

```

QY 1 ATGGGCGCCGCGCGCATCTGCGCGCGAGAAAGCTGGAGAGAGATCCGC 60
DB 1 MetAlaAlaArgAlaSerIleLeuArgIyGluYLeuAspIyTrpGluYsIleArg 20
QY 61 CTGGCGCCGCGCGCGAGAAAGCACTACAGCTGAGACCTGTGTGGCCAGCGCGAG 120
DB 21 LeuArgProGluYglYLeuYsHsIyTrpMetLeuYsHsIleValITrpaIAserTrgIu 40
QY 121 CTGAGAGGCTTGGCCCTGAACCCGCGCTGTGAGAACCGCGAGGGGCTGCAAGCAGATC 180
DB 41 LeuGluArgPheAlaLeuAsnProGluYLeuLeuGluTrpSerGluYsIySylsIle 60
QY 181 ATGAAGCACTGCAAGCCGCGCTGCAAGCCGCGAGAGAGCTGCGACGCTTACAC 240
DB 61 MetYsGluLeuGluInProAlaLeuGluInThrGluYThrGluYLeuYsSerLeuYTrpAsn 80
QY 241 ACCGTCGCGCACTTCTGCTGCTGCGACGCGCGCATGAGAGTCCGCGACCAAGAGGCC 300
DB 81 ThrValAlaThrLeuYTrCySvalHsGluYsIleGluValArgAspThrIySgluAla 100
QY 301 CTGCAACAAGATCGAGAGAGAGCAAGCAAGTCCAGCAGAGAGCCAGAGCGCAAGAG 360
DB 101 LeuAspIyIleGluGluInGluInLeuYsCySgluGluInYsThrGluInAlaYsAla 120
QY 361 GCGGACGCGAGGTGAGCCAGAACTACCCCATGTGACAGACCTGCGAGGGCCAGATGTG 420
DB 121 AlaAspGluYLeuValSerGluInAsnYrProIleValGluInAsnLeuGluInGluInMetVal 140
QY 421 CACCGAGCCATACAGCCCGCGCACCTTGAAACGCTGGGTGTAAGTATTCAGAGAGAGCC 480
DB 141 HsGluInAlaIleSerTrpArgThrLeuAsnAlaITrpaValYValIleGluInIuYsAla 160
QY 481 TTCAGCGCCGAGGTGATCCCATGTTTCAAGCCCTGAGAGAGAGGCGCCAGCCCGAGAGC 540
DB 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluYAlaIThrProGluInAsp 180
QY 541 CTGAACACGATTTGAACAACGCTGGCGGCGCACAGCGCGCCAGTGCATGTGTAAGAGC 600
DB 181 LeuAsnThrMetLeuAsnThrValGluYIleGluInAlaIleMetGluInMetLeuYsAsp 200
QY 601 ACCATCAACGAGAGAGCCCGCGAGTGGAGCCGCTGCAACCCGCTGAGAGCCGCGCGCG 660
DB 201 ThrIleAsnGluGluInAlaIleGluITrpaAspArgLeuHsITrpaValIleGluYProIle 220
QY 661 GCCCGCGCGAGATGCGAGCCCGCGCGAGAGAGCATGCGCGCGCGCCAGCAGCCCTG 720
DB 221 AlaProGluYIleMetArgIuProArgIySerAspIleAlaGluYThrThrSerThrLeu 240
QY 721 CAGAGACAGATGCTGCTGATGACAGCAACCCCGCTGCGCTGGGCGGACATTTACAG 780
DB 241 GlnGluGluInIleAlaITrMetThrSerAsnProProIleProValGluYsPleITryYs 260
QY 781 CGGTGATGATCTGGGCGCTGAACAAAGATCGTGGATTAAGCCCGCTGAGCATCTTACAG 840
DB 261 ArgTrpIleIleLeuGluYLeuAsnYsIleValArgMetYsSerTrpValSerIleLeu 280
QY 841 GACATCCGCGAGAGCCCGCAAGAGAGCCCTTCCGCACTAGTGAAGCGCTTCTTACAGCC 900
DB 281 AspIleArgGluInGluYProYsGluTrpPheAlaArgAspTrpValAspArgPheLeuYsThr 300
QY 901 CTGCGCGCGAGAGAGCCAGCAAGAGAGTGAAGAACTGATGACCGAGACCTGTGCTGTG 960

```

```

DB 301 LeuArgAlaGluInAlaIThrGluInValYsAsnITrMetThrAspIThrLeuLeuVal 320
QY 961 CAGAACGCGCAACCCCGATGGAAGACCATCTCTGGCGGCTTCGCGCCGCGCGCGCACCTG 1020
DB 321 GlnAsnAlaAsnProAspCyAlaYsThrIleLeuArgAlaYsGluYProGluYAlaIThrLeu 340
QY 1021 GAGAGATGATGACCGCTTGCAGAGGCGCTGGGCGCGCGCCCGCGCCAGAGCCCGCTGTG 1080
DB 341 GlnGluMetMetThrAlaCySgluInGluYAlaGluYProGluYsHsIleValITrpaValLeu 360
QY 1081 GCGGAGCGATGAGCGCGAGCCAGCAAGCGCTGAACATCATGATGCAAGAGAGCAACTTCAAG 1140
DB 361 AlaGluAlaMetSerGluInThrAsnSerGluYsHsIleMetMetGluITrGSerAsnPheYs 380
QY 1141 GCGCGCGCGCGCAACGTCATGCTTCACTGCGCGCAAGAGGCGCATATGCGCAAGAGAC 1200
DB 381 GlyProArgArgIleValYsCySvalHsGluYsIleGluYsHsIleValITrpaValLeu 400
QY 1201 TGCGCGCGCCCGCGCAAGAGGCTGTGAGAGTGGCGCAAGAGAGCCAGATGAG 1260
DB 401 CySargAlaProArgIyGluYsGluYCySvalHsGluYsIleGluYsHsIleValITrpaValLeu 420
QY 1261 GACTGCACCTGAGCGCGCAAGGCAACTTCTGGCGCAAGATCTGCGCGCAAGAGAGGCGCGC 1320
DB 421 AspCySvalHsGluYsGluInAlaAsnPheLeuGluYsIleITrProSerHsIleYsGluYArg 440
QY 1321 CCGCGCACTTCTGCGCAAGAGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 441 ProGluAsnPheLeuGluInAsnArgProGluYsIleITrProHsIleValITrpaValLeu 454
QY 1381 CCGCGCGAGAGCTTCTGCGAGAGAGCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB 455 ProAlaGluSerPheArgPheGluGluIThrIThrProAlaProYsGluInGluYProIleGlu 474
QY 1441 GCGGAGCCCTTACCGGAGCCCTTGAAGCGCGCTGCGAGAGCTGTGCGCGAGCGCGCGCTG 1500
DB 475 -----ArgGluProLeuThrSerLeuYsSerLeuPheGluYSerAspProLeu 490
QY 1501 AGCGAG 1506
DB 491 SerGlu 492

RESULT 6
US-10-929-234-5
; Sequence 5, Application US/10929234
; Publication No. US20050123555A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Ketch, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swansstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/10/929,234
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```



OTHER INFORMATION: Description of Artificial Sequence; Note =  
 OTHER INFORMATION: synthetic construct  
 US-10-929-234-5

Alignment Scores:  
 Pred. No.: 2,74e-118 Length: 492  
 Score: 2457.00 Matches: 465  
 Percent Similarity: 95.22% Conservative: 13  
 Best Local Similarity: 92.63% Mismatches: 14  
 Query Match: 86.51% Indels: 10  
 DB: 18 Gaps: 2

US-09-475-704A-4 (1-1509) x US-10-929-234-5 (1-492)

```

QY 1 ATGGGCGCCGCGCCGAGCATCTGCGCGGAGAGACCTGGACAAGTGGAGAAATCCGC 60
   |||
   |||
   |||
Db 1 MetAlAlARglaSerlleuARgIyGluLyLeuAspIyStrpGluysIleARg 20
   |||
   |||
   |||
QY 61 CTGGCGCCGCGCGGCAAGAAACAATACTGCTGAACCACTGCTGGCGCGAGCCGCGAG 120
   |||
   |||
   |||
Db 21 LeuARgPrOgIyGluLyshIstIyMeLyLeuLyshIleValItrpAlaserARgIu 40
   |||
   |||
   |||
QY 121 CTGAGAGGCTTCGCTGAAACCCGCGCTGCTGAGACCGCGAGGGCTGCAAGCAGATC 180
   |||
   |||
   |||
Db 41 LeuGIuARgPheAlaleuAsnPrOgIyLeuLeuGIuThrSerGIuIyCyshIle 60
   |||
   |||
   |||
QY 181 ATGAAACAGCTGACGCGCGCTGACAGACCGCGAGACCGAGAGCTGCGACCTGTACAAC 240
   |||
   |||
   |||
Db 61 MeLyGIuLeuGIuPrOAlaleuGIuThrGIuThrGIuLeuLySerIeuTyran 80
   |||
   |||
   |||
QY 241 ACCGTGGCCACCTGTACTGCTGACAGCGCGGATGAGGTCCGCGACACCAAGAGGCC 300
   |||
   |||
   |||
Db 81 ThrValAlatrIeuTyCyshValhIsgIuLyIleGIuValARgshItrpIySGluAla 100
   |||
   |||
   |||
QY 301 CTGCAACAAGTCAGAGAGAGACAGAAACAAGTCCCAAGACAGAGAGCGCGAGAGAG 360
   |||
   |||
   |||
Db 101 LeuAspLyIleGIuGIuGIuGIuAsnLyCyshGIuGIuThrGIuGIuAla 120
   |||
   |||
   |||
QY 361 GCGGAGCGCAAGGTGAGCAAACTACCCCATGTGCAAGACCTGCGAGGCGCAGATGTG 420
   |||
   |||
   |||
Db 121 AlaAspGIyLeuValSerGIuAsnTyPrOIlleValGIuAsnLeuGIuGIuMetVal 140
   |||
   |||
   |||
QY 421 CACGAGGCATCAGCCCGCGCACCTGAAACGCTGGGTGAAGTATCGAAGAAAGGCC 480
   |||
   |||
   |||
Db 141 HIsGIuAlaIleSerPrOARgThrIleuAsnAlaItrpValIyValIleGIuGIuAla 160
   |||
   |||
   |||
QY 481 TTCAGCGCGAGGTGATCCCATGTTCACCGCGCTGAGGAGGCGCGACCCCGCGAGAC 540
   |||
   |||
   |||
Db 161 PheSerPrOgIuValIlePrOmetPheThrAlaleuSerGIuGIuAlaThrPrOgIuAsp 180
   |||
   |||
   |||
QY 541 CTGAACACAGATTTGAACAACCGTGGCGGCGACACAGCGCGCATGCAATGTGAAGAC 600
   |||
   |||
   |||
Db 181 LeuAsnThrMetLeuAsnThrValGIuGIuHIsGIuAlaAlaMetGIuMetLeuysAsp 200
   |||
   |||
   |||
QY 601 ACCATCAAGAGAGGCGCGCGAGTGGACCGCTTCACCCCGCTGAGGCGGCGCGCTG 660
   |||
   |||
   |||
Db 201 ThrIleAsnGIuGIuAlaIleGIuTrpAspARgIuAsnIlePrOValhIsgIyPrOIlle 220
   |||
   |||
   |||
QY 661 GCGCGCGCGAGATGCGCGACCCCGCGGAGAGACATGCGCGGCGCGCACCGACCCG 720
   |||
   |||
   |||
Db 221 AlaPrOgIyGIuMetARgIuPrOARgGIySerAspIleAlaGIyThrThrSerThrLeu 240
   |||
   |||
   |||
QY 721 CAGAGACAGATCGCTGATGACACAGAACCCCGCGTGGCGCGCGCATTTACAAG 780
   |||
   |||
   |||
Db 241 GIuGIuGIuHIsIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 260
   |||
   |||
   |||
QY 781 CGGTGATCATCTGGGCGCTGAACAAGATGTGCGATGTACAGCCCGCTGAGATCTTG 840
   |||
   |||
   |||
Db 261 ARgItrpIleIleuGIuLeuAsnLyIleValARgMetTySerPrOValSerIleLeu 280
   |||
   |||
   |||
QY 841 GACATCGGAGGCGCGCAAGAGGCGCTTCGCGGATAGTGAAGCGCTTTCAGAGAC 900
   |||
   |||
   |||
Db 281 AspIleARgGIuGIuPrOySGIuPrOARgAspIyTrValAspARgPhePheLyStr 300

```

```

QY 901 CTGGCGCGCGAGACAGACCCCAAGAGAGTGAAGAACTGATGACCGGACACCTGTGCTG 960
   |||
   |||
   |||
Db 301 LeuARgAlaGIuGIuAlaIleThrGIuGIuValIyAsnItrpMetThrAspThrIleuVal 320
   |||
   |||
   |||
QY 961 CAGAACGCGCAACCCCGATGCAAGACCAATCTGCGCGCTGCGCGCGCGCGCGCGCGCTG 1020
   |||
   |||
   |||
Db 321 GIuAsnAlaAsnPrOAspCyshIyThrIleuARgAlaleuGIyPrOgIyAlaIleu 340
   |||
   |||
   |||
QY 1021 GAGAGATGATGACCGCTGCCAGGCGGTGGCGGCGCGCGCGCGCGCGCGCGCGCTG 1080
   |||
   |||
   |||
Db 341 GIuGIuMetMetThrAlaCyseGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 360
   |||
   |||
   |||
QY 1081 GCGGAGCGGATGACCGACCGCAACAGCGGTGAACATCATGATGACAGAAAGCACTTACA 1140
   |||
   |||
   |||
Db 361 AlagIuAlaMetSerGIuThrAsnSerGIyAsnIleMetMetGIuARgSerAsnPheLy 380
   |||
   |||
   |||
QY 1141 GCGCGCGCGCGCAAGTCAAGTCTCAACTGCGCGCAAGAGGCGCGACATGCGCAAGAAC 1200
   |||
   |||
   |||
Db 381 GIyPrOARgARgIleValIyCyshPheAsnCyshIySGIuGIuHIsIleAlaARgAsn 400
   |||
   |||
   |||
QY 1201 TCCCGCGCGCGCGCAAGAGGCTGCTGAAGTGCAGAGAGAGGCGCGACAGATGAAG 1260
   |||
   |||
   |||
Db 401 CyshARgAlaPrOARgLyshIySerIyPrOyCyshIySGIuGIuGIuHIsGIuMetLy 420
   |||
   |||
   |||
QY 1261 GACTGCACCGAGCGCGCAAGCACTTCTGCGCAAGATCTGCGCAAGCGCGCGCGCGC 1320
   |||
   |||
   |||
Db 421 AspCyshThrGIuARgIuAlaAsnPheLeuGIySIleItrpPrOserIleysGIyARg 440
   |||
   |||
   |||
QY 1321 CCGCGCACTTCCTGCAAGACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380
   |||
   |||
   |||
Db 441 PrOgIyAsnPheLeuGIuAsnARgPrOgIu-----PrOThrAlaPrO 454
   |||
   |||
   |||
QY 1381 CCGCGCGAGAGCTTCGCTGAGAGAGACCAACCCCGCGCGCAAGAGCGCGCGCGCGC 1440
   |||
   |||
   |||
Db 455 PrOAlaGIuSerPheARgPheGIuGIuThrItrpPrOAlaPrOyGIuGIuPrOIlleGIu 474
   |||
   |||
   |||
QY 1441 CCGGAGCCCTTACCGGAGCGCGCTGACCGCGCTGCGAGCGCTGTGCGAGCGCGCGCTG 1500
   |||
   |||
   |||
Db 475 -----ARgIuPrOleuThrSerIeuLySerIeuPheGIySerAspPrOleu 490
   |||
   |||
   |||
QY 1501 AGCGAG 1506
   |||
   |||
   |||
Db 491 SerGIu 492
   |||
   |||
   |||

```

RESULT 7  
 US-10-332-413-8  
 Sequence 8, Application US/10332413  
 Publication No. US20040116660A1  
 GENERAL INFORMATION:  
 APPLICANT: Johnstone, Robert Edward  
 APPLICANT: Swanson, Ronald Ivar  
 APPLICANT: Morris, Lynn  
 APPLICANT: Karim, Salim Abdool  
 APPLICANT: Williamson, Carolyn  
 TITLE OF INVENTION: Processes for the Selection of HIV-1 Subtype C Isolates, Selected  
 TITLE OF INVENTION: Subtype Isolates, Their Genes and Modifications and Derivatives  
 FILE REFERENCE: 45669-281993  
 CURRENT APPLICATION NUMBER: US/10/332,413  
 CURRENT FILING DATE: 2003-09-12  
 PRIOR APPLICATION NUMBER: PCT/IB01/01208  
 PRIOR FILING DATE: 2001-07-09  
 PRIOR APPLICATION NUMBER: US 60/216,995  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: ZA 2000/3437  
 PRIOR FILING DATE: 2000-07-10  
 PRIOR APPLICATION NUMBER: ZA 2000/4924  
 PRIOR FILING DATE: 2000-09-15  
 NUMBER OF SEQ ID NOS: 32  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 8  
 LENGTH: 631  
 TYPE: PRT

ORGANISM: Human immunodeficiency virus type 1  
US-10-332-413-8

## Alignment Scores:

Pred. No.:	2,76e-118	Length:	631
Score:	2457.00	Matches:	465
Percent Similarity:	95.22%	Conservative:	13
Best Local Similarity:	92.63%	Mismatches:	14
Query Match:	86.51%	Indels:	10
DB:	16	Gaps:	2

US-09-475-704A-4 (1-1509) x US-10-332-413-8 (1-631)

```
QY 1 ATGGGCGCGCGCGCATCTGCGCGCGAGAGACTGTGACAAAGTGGAGAAATCCGC
DB 31 MetAlaAlaArgAlaSerIleLeuArgGlyGluIuLeuAapIyStrpGluIyStleArg
QY 61 CTGGCGCGCGCGCGCAAGAACAATCATGTGTGAACACTGGTGTGGCCAGCCGCGAG
DB 51 LeuArgProGlyGlyIuIyIyShIyStryMetLeuIyHleIleValItrpAlaSerArgIu
QY 121 CTGGAGGGGTTGGCTTGAACCCCGGCTGTGTGAGAACCGCGAGGGCTGTGACAGATC
DB 71 LeuGluArgPheAlaLeuAanProGlyLeuLeuGluThrSerGluIyCybIySgIuIle
QY 181 ATGAAGACAGTCAAGCGCGCGCTGCAGACCGGCAAGAGAGCTGGCGAGCTGTACAC
DB 91 MetIySgIuIeGluInProAlaLeuGluInThrGluIuLeuIySerIeutyIyAan
QY 241 ACCGTGGCCACCTGTACTGTCTGTGCAGCGCGCGCATGAGGTCCGCGACCAAGAGGCC
DB 111 ThrValAlaIleThrLeuIyCybValHISglIuIyIleGluValArgPheThrIySgIuAla
QY 301 CTGGCAAAATGTGAGAGAGAGAGAGAACAGTCCCGACAGAGACCCAGAGGCCAAGAG
DB 131 LeuAapIyIleGluIuIeGluInIuIeAanIyCybGluInIyThrGluInAlaIySgIuAla
QY 361 GCCGACGCGAGGTGAGCGCAAACTACCCCATGTGTGCAGAACCTGCAGAGGCCAGATG
DB 151 AlaAapGlyIyValIeSerIeIyAanIyProIleValGluInIeGluInIyMetVal
QY 421 CACCAAGGCATCAAGCCCGCGACCTGAAAGCCTGGGTGAAGGTATGAGAGAGAGGCC
DB 171 HisGluAlaIleSerProArgThrLeuAanAlaItrpValIleGluIuIySgIuAla
QY 481 TTCAGCGCGGAGTGTATCCCATGTTTCAACCGCTGAGAGAGAGGCCACCCCGAGAGC
DB 191 PheSerProGluValIleProMetPheThrAlaLeuSerGluIyAlaIthrProGluAap
QY 541 CTGAACACGATGTTGAACACCGTGGGCGGCGACAGCGCGCCATGAGATGTGTAAGAGC
DB 211 LeuAanThrMetLeuAanThrValIeGlyIyHISgluInAlaIleMetIeLeuIyAap
QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGAGCGCGCTGCACCCCGGTGAGCGCGCGCGCTG
DB 231 ThrIleAanGluIuAlaIleGluIuItrpAapArgLeuHISProValHisAlaGlyProIle
QY 661 GCGCGCGCGCGCATGGCGCAACCCCGCGCGACAGACATGCGCGCGCGCCACAGCACCTG
DB 251 AlaProGlyIyGluMetArgIuIeProArgIySerAspIleAlaGlyThrIySerThrIe
QY 721 CAGAGACAGATGCGCTGATGACAGCAACCCCGCGCGCGCGCGGTGGGCGAGATCTTACAG
DB 271 GlnGluGluInIleAlaItrpMetIySerAanProProIleProValGlyAapIleIyIyIyS
QY 781 CGGTGATCATCTGTGGCGCTGAAACAAGATCGTGGAGATGACGCCCGCTGACATCTCTG
DB 291 ArgTrrpIleIleIeGluIyLeuAanIySISleValArgMetIySerProIleSerIleIe
QY 841 GACATTCGCGAGAGCGCGCGCAAGAGCGCTTCCGCGACTAGTGAACCGCTTTTCAAGACC
DB 311 AapIleArgGluIyProIySgIuProPheArgAapIyValAaAapArgPheAapIyStr
330
```

```
QY 901 CTGGCGCGCGAGAGAGCGCGCGCAAGAGACTGTGAAGAACTGGATGACCGAGACCTGTGGTG
DB 331 LeuArgAlaGluIuAlaIleThrGluInIuIySgIuAanIyMetIyThrAapThrIleuVal
QY 961 CAGAAAGCCAAACCCGAGTGCAGAGACATCTGTGGCGCTGTGGCGCGCGCGCGCACCTG
DB 351 GlnAanAlaAanProAapCybIyStrIleLeuArgAlaIeGlyProGlyAlaIthrIe
QY 1021 GAGGAGATGATGACCGCGCGTGCAGGCGCGTGGCGCGCGCGCGCGCGCAAGCGCGTGTG
DB 371 GlnGluMetMetThrAlaCybGluIyValIeGlyIyProGlyIyHISgluIyAlaIyValIe
QY 1081 GCCGAGCGGATGAGCGCGCGCAAGAGCGCGTGAACATCATGATGACAGAAAGAACTTCAAG
DB 391 AlaGluAlaIeMetSerGluInThrAanSerGlyAanIleMetMetGluArgSerAanPheIyS
QY 1141 GCGCGCGCGCGCAAGCTCAAGTCTTCAACTGTGGCGCAAGAGAGCGCGACATGCGCAAGAC
DB 411 GlyProArgArgIleValIyIeCybPheAanCybGlyIySgIuIyHISgluIyAlaIyAan
QY 1201 TGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCGCAAGAGAGCGCGCACCAATGAAG
DB 431 CybArgAlaProArgIySgIyCybIyIyCybIyIySgIySgIyHISgluIyHISgluIyS
QY 1261 GACTGACACCGAGCGCGCGCAAGCTTCTGTGGCGCAAGATCTGCGCGCGCACAGAGCGCGC
DB 451 AapCybIyThrGluIyGluAlaAanPheLeuIyIySISleIyIyProIleHISgluIyArg
QY 1321 CCGGCGAACTTCTGTGACAGACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC
DB 471 ProGluAanPheLeuGluInAanArgProGlu-----ProThrAlaPro
QY 1381 CCGCGCGAGAGTTCGTTGAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGC
DB 484 471 ProGluAanPheLeuGluInAanArgProGlu-----ProThrAlaPro
QY 1441 CGCGAGCCCTTACCGGAGAGCGCGCGTGCAGCGCGCGTGTGCGAGCGCGCGCGCTG
DB 504 1441 CGCGAGCCCTTACCGGAGAGCGCGCGTGCAGCGCGCGTGTGCGAGCGCGCGCGCTG
QY 1501 AGCCAG 1506
DB 521 Sergin 522
RESULT 8
US-10-780-507-98
; Sequence 98, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MILLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780, 507
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial sequence
```



```

; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Most recent common ancestor reconstruction of clade C gag protei
; US-10-780-507-96

Alignment Scores:
  Pred. No.:      6,28e-118      Length:      492
  Score:          2450.00        Matches:      463
  Percent Similarity: 95.22%      Conservative: 15
  Best Local Similarity: 92.23%      Mismatches:  14
  Query Match:      86.27%        Indels:       10
  DB:               18           Gaps:         2

US-09-475-704a-4 (1-1509) x US-10-780-507-96 (1-492)

QY      1 ATGGGCGCGCGCGCGCATCTCGCGCGCGAGAAAGCTGAGCAAGTGGAGAAATCCGC 60
DB      1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAbpThrTrpGluysIleArg 20
QY      61 CTGGCGCGCGCGCGCGAGAAAGCACTACCTGAGCAAGCACTGCTGGCGCGCGCGAG 120
DB      21 LeuArgProGlyGlyLeuysHlyyrmcIlelyHnIleuValTrrAlaSerArgGlu 40
QY      121 CTGAGGCGCTTGGCGCGCGCGCGCGCTGAGAGACCGCGAGGCGCTGAGAGAGATC 180
DB      41 LeuGluArgPheAlaLeuAbnProGlyLeuLeuGluThrSerGluGlySerGlnIle 60
QY      181 ATGAAGACAGCTGAGCGCGCGCGCTGAGAGACCGCGAGAGCTGCGACCTGTACAC 240
DB      61 IlelyGlnIleuGlnProAlaLeuGlnThrGlyThrGluIleuysSerLeuYrAen 80
QY      241 ACCGCGCGCGCGCGCTGTACTGTGCGCGCGCGCGAGTGGATCGCGCGAGCAAGAGCGC 300
DB      81 ThrValAlaThrLeuYrCyvalHisGlnArgIleGlnValArgAspThrlysgIuAla 100
QY      301 CTGAGCAGAGATGAGAGAGAGAGAGAAACAAGTCCAGCAGAGAACCGAGCGCGCAAGAG 360
DB      101 LeuAbpIyIleGluGluGlnIleuAbnysSerGlnGlnIleuThrGlnIleuAlaGluAla 120
QY      361 GCGGAGCGGAGAGAGAGAGAGAACTACCCCATCTGTGAGAGACCTTGAGGCGCGAGATGTG 420
DB      121 AlaAbpGlyLyvalSerGlnAbnYrProIleValGlnAbnIleuGlnIyGlnMetVal 140
QY      421 GACCGAGCGCATGAGCGCGCGCGCGCGCGCGCGCGTGTGAGTGAAGTGAAGAGAGCGC 480
DB      141 HisGlnAlaIleSerProArgThrLeuAbnAlaTrrValIleGluGluIyValAla 160
QY      481 TTTCAGCGCGAGGTGATCCCATGTTTCACCGCGCTGAGAGAGAGAGCGCGCGCGAGAGAC 540
DB      161 PheSerProGluValIleProMetPheThrAlaLeuSerGlnGlyAlaThrProGlnAbp 180
QY      541 CTGAACACAGATGTTGAACACCGTGGCGCGCGCACAGCGCGCGCATGCAATGCTTAAGAGAC 600
DB      181 LeuAbnThrMetLeuAbnThrValGlyGlyHisGlnAlaAlaMetClnMetLeuysAbp 200
QY      601 ACCATCAACAGAGAGCGCGCGAGTGGAGCGCGCGTGCACCGCGTGACGCGCGCGCGCGGTG 660
DB      201 ThrIleAbnGluGluAlaIleGluTrrPArArgLeuHisAbpProValHisAlaGlyProVal 220
QY      661 GCGCGCGCGAGATGCGCGCGCGCGCGCGAGCATGCGCGCGCGCGCGCGCGCGCGCGCTG 720
DB      221 AlaProGlyGlnMetArgIleuProArgIySerAbpIleAlaGlyThrThrSerThrIleu 240
QY      721 CAGAGACAGATGCGCTGTGATGACAGAGAACCGCGCGCGTGGCGCGCGAGCATTTACAG 780
DB      241 GlnGluGlnIleAlaIleThrMetThrSerAbnProProIleProValGlyAspIleYrIyys 260
QY      781 CGGTGATCATCTGGCGCGCGTGAACAGATCGCGAGTGAAGCGCGCGCGCGCGCGCGCTG 840
DB      261 ArgTrrIleIleuGlnIleuAbnIySIIleValArgMetYrSerProValSerIleu 280

```

```

QY      841 GACATCCGCGAGGCGCGCGAGAGCGCTTCCGCGCATGATGAGACCGCTTCTCAAGACC 900
DB      281 AspIlelysgIlnGlyProIyGluProPheArgAspYrValAlaBpArgPhePheIythr 300
QY      901 CTGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB      301 LeuArgAlaGluGlnAlaThrGlnAbpValIyysAbnTrrMetThrAspThrIleuVal 320
QY      961 CAGAGCGCGAACCGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB      321 GlnAbnAlaAbnProAspCyAbpThrIleuAbnAlaLeuGlyProGlyAlaThrIleu 340
QY      1021 GAGAGATGATGAGCGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB      341 GlnGluMetMetThrAlaCyGlnGlyValGlyIyProSerHnIlelyAlaArgValIleu 360
QY      1081 GCGGAGGAGATGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB      361 AlaGluAlaMetSerGlnAlaAbnAbnThrAbnIleMetMetGlnArgGlyAsnPhelys 380
QY      1141 GCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB      381 GlyProArgArgIleValIyCyPheAbnYrGlyGlyGlyGlnIyHisIleAlaArgAsn 400
QY      1201 TGCCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB      401 CysArgAlaProArgIyGlyGlyCyThrIyPheGlyGlyGlyGlnIyHisIleGlnMetIyys 420
QY      1261 GACTGACCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB      421 AspCyThrGluArgGlnAlaAbnPhelenuIyIySIIeTrrProSerHnIlelyGlyArg 440
QY      1321 CCGCGCACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB      441 ProGlyAbnPhelenuIleGlnSerArgProGlu-----ProThrAlaPro 454
QY      1381 CCGCGCGAGAGCTTCCGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB      455 ProAlaGlnSerPheArgPheGluGluThrThrProAlaProIyGlnGluProIyysAbp 474
QY      1441 GCGGAGCGCTTACCGGAGAGCGCGCTGACCGCGCTTGGCGAGCGCTTTCGGAGCGCGCGCTG 1500
DB      475 -----ArgGluProLeuThrSerIleuysSerIleuPheGlySerAbpProIeu 490
QY      1501 AGCGAG 1506
DB      491 SerGln 492

RESULT 10
; US-10-339-217-109
; Sequence 109, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Mengang
; TITLE OF INVENTION: Methods for Identifying Compounds which Inhibit Binding of
; TITLE OF INVENTION: Nucleosidesd 7 Protein to HIV-1 RNA
; FILE REFERENCE: 111021.143 (ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 492
; TYPE: PRT
; ORGANISM: HIV-1
; US-10-339-217-109

```

## Alignment Scores:

Pred. No.: 2,05e-117 Length: 492  
 Score: 2440.00 Matches: 464  
 Percent Similarity: 94.42% Conservative: 10  
 Best Local Similarity: 92.43% Mismatches: 18  
 Query Match: 85.92% Indels: 10  
 DB: 14 Gaps: 2

US-09-475-704a-4 (1-1509) x US-10-339-217-109 (1-492)

```

QY      1 ATGGGCGCCGCGCCGAGCATCTGCGCGCGAGAACCTGAGCAAGTGGAGAAATCCGC 60
DB      1 MetGlyAlaSerAlaSerIleLeuArgGlyGlyLysLeuAspLysTrpGluLysIleArg 20

QY      61 CTGGCGCCCGCGCGCAAGAACATTCATGCTGAGACACCTGGTGGCCGACGCCGCGAG 120
DB      21 LeuArgProGlyGlyLysLysLysTrpArgLeuLysHisLeuValTrpAlaSerArgGlu 40

QY      121 CTGGAGGGGCTTGGCCCTGAACCCCGGCTGAGAGACCGCGAGGGGCTGCAAGGATC 180
DB      41 LeuGluArgPheAlaLeuAsnSerGlyLeuLeuGluThrAlaGluGlyCysLysGlnIle 60

QY      181 ATGAAGCAAGCTGCAAGCCCGCTGCAAGACCGGACCGAGAGCTGCGACCTGTACAAC 240
DB      61 IleLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTrpAsn 80

QY      241 ACCGTGGCCACCTGTACTGCTGCGTGCACGCGCGGATCAGAGTCCGCGACACCAAGAGGCC 300
DB      81 ThrValAlaThrLeuLysCysValHisAlaGlyIleGluValArgAspThrLysGluAla 100

QY      301 CTGGCAAGATGAGGAGGAGGAGGAGCAAGATGCCCGAGAGAGACCCGAGGCGCAAGAG 360
DB      101 LeuAspLysIleGluGluGlnGlnAsnLysCysGlnGlnLysThrGlnAlaLysGln 120

QY      361 GCGGACGCGAGGTGAGCGCAAGACTACCCCATGTCGACAGACCTGCGAGGCGCGAGTGTG 420
DB      121 AlaAspGlyLysValSerGlnAsnTrpProIleValGlnAsnLeuGlnIleGlnMetVal 140

QY      421 CACGAGGCGCATCAGCCCGCGACACCTGAACGCGCTGGGTGAGGTGATCGAGAGAGGCC 480
DB      141 HisGlnProIleSerProArgThrLeuAsnAlaTrpValIleGluGluLysAla 160

QY      481 TTCAGCGCCGAGGTATCCCGCATGTTCAAGCCGCTGAGAGGAGGCGCGCCACCCCGAGAC 540
DB      161 PheSerProGluValIleProMetPheThrAlaLeuSerGlnGlyAlaThrProGlnAsp 180

QY      541 CTGAACACGATGTTGAACACCGTGGGCGGCGCACAGGCGCGCATGAGATGCTGAAGAC 600
DB      181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaHisMetGlnMetLeuLysAsp 200

QY      601 ACCATCAACGAGAGGCGCGCGAGTGGACCGCTTGACACCCCGTGCAGGCGCGCCCGT 660
DB      201 ThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisProValHisAlaGlyProVal 220

QY      661 GCGCCCGCGGACATGGCGCACCCCGCGGACGAGACATGGCGGCGCGCACAGACCTTG 720
DB      221 AlaProGlyGlnMetArgGlnProAlaGlySerAspIleAlaGlyThrSerThrLeu 240

QY      721 CAGGACGATGCGCTGATGACACAGCAACCCCGCTGCGCGGCGCGACATCTTCAAG 780
DB      241 GlnGluGlnIleThrTrpMetThrSerAsnProProValProValGlyAspIleTrpLys 260

QY      781 CGGTGATCATCTTGGGCTTGAACAAAGATGTGCGGATGTACAGCCCGGTGAGCATCTG 840
DB      261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProValSerIleLeu 280

QY      841 GACATCGCGAGGCGCGCAAGAGCCCTTCCGCGATAGTGAACCGCTTCAAGAC 900
DB      281 AspIleLysGlnGlyProLysGlnProPheArgAspTrpValAspArgPhePheLysVal 300

QY      901 CTGGCGCGCGAGCGCGCACCGCAGACGTGAAGAACTGATGACCGAGACCTTGCTGGTG 960
DB      301 LeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeuVal 320
  
```

```

QY      961 CAGAGCGCAACCCCGAGTGAAGACATCTCTGCGCGCTTGGGCCCGCGCGCACCTG 1020
DB      321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProAlaIleSerLeu 340

QY      1021 GAGAGATGATGACCGCTGCGCGAGGCGTGGGCGGCGCCGCGCACAGGCGCGCTG 1080
DB      341 GlnGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLeuAlaArgValLeu 360

QY      1081 GCGGAGCGATGAGCGCAGCGCAACAGCGTGAACATCATGATGACAGAGCAACTTCAAG 1140
DB      361 AlaGluAlaMetSerGlnAlaAsnThrThrAsnIleMetMetGlnLysSerAsnPheLys 380

QY      1141 GCGCCCGCGCGCGCAAGCTGAAGTCTTCAACTGCGCGCAAGAGGCGCAATCGCCCAAGAC 1200
DB      381 GlyProArgArgThrValLysCysPheAsnLysCysGlyLysGlnGlyHisIleAlaLysAsn 400

QY      1201 TGCGCGCGCGCGCGCAAGAGGCTGTGGAAGTCCGCGCAAGAGGCGCACAGATGAAG 1260
DB      401 CysArgAlaProArgLysLysGlyGlyCysTrpLysCysGlyLysGlnGlyHisGlnMetLys 420

QY      1261 GACTGACACCGAGCGCGCAGCGCAACTTCTGGGCAAGATCTGCGCGCACAGGCGCGC 1320
DB      421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440

QY      1321 CCGCGCAACTTCTGCAAGAACCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB      441 ProGluAsnPheLeuGlnSerArgProGlu-----ProThrLysPro 454

QY      1381 CCGCGCGAGGCTTCCGCTGAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB      455 ProAlaGluSerPheArgPheGluGlnThrThrProAlaProLysGlnGluProLysAsp 474

QY      1441 CGCGAGCCCTTACCGGAGCCCGCTGACCGCGCTGCGGACCTGTTGGAGGAGCCCTG 1500
DB      475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490

QY      1501 AGCCAG 1506
DB      491 SerGln 492

RESULT 11
US-10-339-217-110
; Sequence 110, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Maior, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds which Inhibit Binding of
; FILE REFERENCE: 111021,143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 492
; TYPE: PRN
; ORGANISM: HIV-1
US-10-339-217-110

Alignment Scores:
Pred. No.: 6.68e-117 Length: 492
Score: 2430.00 Matches: 458
Percent Similarity: 94.82% Conservative: 18
Best Local Similarity: 91.24% Mismatches: 16
Query Match: 85.56% Indels: 10
DB: 14 Gaps: 2
  
```

US-09-475-704A-4 (1-1509) x US-10-339-217-110 (1-492)

```

QY 1 ATGGGGCGCCGGCCGAGCATCTCTGGCGGCGAGAACTGGAACAAGTGGAGAAATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyLysLeuAspThrTrpGlyLysIleArg 20
QY 61 CTGGCGCCCGCGCGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGCCAGCGCGAG 120
DB 21 LeuArgProGlyGlyLysIleValMetLeuLysIleValTrpAlaSerArgGly 40
QY 121 CTGGAGAGGCTTGGCCCTGGAACCCCGGCTGTGGAGACCGCGAGGCTGCAAGACAGATC 180
DB 41 LeuGluArgPheSerLeuAsnProGlyLeuLeuGluThrSerGluGlyLysGlnIle 60
QY 181 ATGAAGACACTGACGCGCCCGCTGACAGACCGGACCGAAGAGCTGGGACCTGTACAC 240
DB 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuThrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCAGCGCGGCGATCGAGGTCCGCGACCAAGAGGCGC 300
DB 81 ThrValAlaThrLeuPheCysValIleGlyLysIleAlaValAlaArgAspThrLysGluAla 100
QY 301 CTGGACAAGATCGAGAGGAGGACGAACAAGTCCGACAGACAGACCCGACGCGCAAGAG 360
DB 101 LeuAspLysIleGluGluGlnLeuAsnLysSerGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCCGACGGGAGAGTGAAGCCAGAACTACCCCATCTGTGCAACAACCTGACGAGGCGCAGATGTC 420
DB 121 AlaAspGlyThrValSerGlnAsnLysrProIleValGlnAsnLeuGlnIleGlnMetVal 140
QY 421 CACGAGGCGATCAGCGCCGCGACCTGGAAGCGCTGGGTGAAGGTATCGAGAGGAAGGCGC 480
DB 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
QY 481 TTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGGCGAGGCGCGCACCCCGCAGAC 540
DB 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTTGAACAACCGTGGGCGGCGACCGCGCCGACATGCAAGTCTGAAGAGC 600
DB 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCGGCTGACCGCCGTGACGCGCGCGCGCCGTC 660
DB 201 ThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisProValHisAlaGlyProIle 220
QY 661 GCCCGCGCGCAAGTGGCGACCCCGCGGCGAGGACATGTCGCGCGCGCGCACGACCTGT 720
DB 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGGAGCAATGCGCTGTGATGACCAAGCAACCCCGCGTCCCGTGGGCGCACTTCAAG 780
DB 241 GlnGluGlnIleLeuAlaTrpMetThrAsnAsnProProValProValGlyAspIleLys 260
QY 781 CGGTGATCATCTGGGCTGGAACAAGATCGTGGGATGTACAGCCCGCTGAGACATCTGT 840
DB 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProValSerIleLeu 280
QY 841 GACATTCGCGCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGACCGCTTCTTCAAGAC 900
DB 281 AspIleArgGlnGlyProLysGluProPheArgAspTrpValAspArgPhePheLysThr 300
QY 901 CTGGCGCGCGAGAGCGCGACCAAGACCTGAAGAACTGATGATGACGAGACCTGTGCTGGT 960
DB 301 LeuArgAlaGluGlnIleThrGlnAspValLysAsnTrpMetThrAspThrLeuVal 320
QY 961 CAGAAAGCCCAACCCCGCAAGACCAATCTGTGCGGCTGTGCGCGCGCGCGCACCTGT 1020
DB 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGGAGATATGACCGCTGTGCCAGGCGGTGGCGGCGCGCGCGCACAGGCGCGCGTGT 1080

```

```

DB 341 GluIleMetMetThrAlaCysGlnGlyValGlyLysProSerHisLysAlaArgValLeu 360
QY 1081 GCCGAGGGGATGAGCGCAGACCAACAGCTGAACATCAAGATGCAAGAAAGCACTTCAAG 1140
DB 361 AlaGluAlaMetSerGlnThrAsnAsnAlaAsnIleMetMetGlnArgSerAsnPheLys 380
QY 1141 GGCGCCCGCGCGCAACGTCAGTGTCTTCAACTGCGGCGCAAGAGGCGCACATCGCCCAAGAC 1200
DB 381 GlyProAlaArgIleIleLysCysPheAsnLysGlyLysGluGlnHisLeuAlaArgAsn 400
QY 1201 TGCCGCGCCCGCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGCGCACCAAGATGAAG 1260
DB 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlnHisGlnMetLys 420
QY 1261 GACTGACACCGAGCGCGCAAGCACTCTGTGGCAAGATCTGGCCGACCAAGGCGCGC 1320
DB 421 AspCysThrGluArgGlnAlaAsnPheLeuLysLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCCGCAACTTCTGTCAACAACCGGACGAGCGCGCGCGCCCGCCACCGGTCCGACGCGCC 1380
DB 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454
QY 1381 CCCGCGAGAGCTTCCGCTTGAAGAGCAACCCCGCGCCCAAGACAGACCCCAAGAGAC 1440
DB 455 ProAlaGlnSerPheArgPheGluGluThrThrProAlaProLysGlnGluProArgGlu 474
QY 1441 CGCGAGCCCTTACCGGAGACCCCTGACCGCCCTGTGGCGAGCCTTGTGGGAGGCGCCCTG 1500
DB 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
DB 491 SerGln 492

```

RESULT 12  
US-10-339-217-140  
Sequence 140, Application US/10339217  
Publication No. US20030198648A1  
GENERAL INFORMATION:  
APPLICANT: Buechter, Douglas  
APPLICANT: Hou, Xiaohong  
APPLICANT: Marlor, Christopher W.  
APPLICANT: Rice, William G.  
APPLICANT: Yang, Wengang  
TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of  
FILE REFERENCE: 111021.143 (ACH-US1)  
CURRENT FILING DATE: 2003-01-09  
PRIOR APPLICATION NUMBER: US 60/347,369  
PRIOR FILING DATE: 2002-01-11  
NUMBER OF SEQ ID NOS: 144  
SOFTWARE: FASTSeq for Windows Version 4.0  
SEQ ID NO 140  
LENGTH: 491  
TYPE: PRF  
ORGANISM: HIV-1  
US-10-339-217-140

Alignment Scores:  
Pred. No.: 1,83e-116 Length: 491  
Score: 2421.50 Matches: 461  
Percent Similarity: 94.82% Conservative: 15  
Best Local Similarity: 91.83% Mismatches: 15  
Query Match: 85.26% Indels: 11  
DB: 14 Gaps: 3

US-09-475-704A-4 (1-1509) x US-10-339-217-140 (1-491)

```

QY 1 ATGGGCGCCCGCGCAGATCTCTGGCGGCGAGAAAGTGGAGAAATCCGC 60
DB 1 MetGlyAlaArgAlaSerIleLeuArgGlyLysLeuAspThrTrpGlyLysIleArg 20

```

OY		61	CTGGCGCCCGGCGGAAGAAGACATCAAGCTGAAGACACTGTGTGGGCCGACCGCGAG	120
Db		21	Leuaagprrgglgilylvayshstfynrmetleuylshlslevualtpralaserahgllu	40
OY		121	CTGAGAGGCCTTCGCCCTTGAAACCCCGGCTGTGAGAACCGCGAAGGCTGCAAGCATC	180
Db		41	LeugluargpnealiaValasnProglyneuLeugluThraligluclyyelysgntile	60
OY		181	ATGAAGCAGCTGCACGCCGCCCTTGCACAGCCGCGCACCGAGAGACTGTGCAGCCTGTACAAC	240
Db		61	IleyysgInteuglnProalaenugInthrnglyThnrgluglmeuhargserleutyraen	80
OY		241	ACCGGGGCAACCCTGTACTGCGGTGACCGCGGCATCGAGSTCCGCGACACCAAGAGGCC	300
Db		81	ThrvalatathLeutyrcysvalhhbaIagluileguvalArgaspthrysgluvala	100
OY		301	CTGCAACAAGATGAGAGAGAGACAGAAACAAGTCCACAGAAAGACCAGAGGCTCAAGAG	360
Db		101	LeuaeparaglieglucglucglumenuyserrgIntuglyThrdlnmlabenglu	120
OY		361	GCCGACGGCAAGGTGAGGCCAAGACTACCCCATTGTGCAGAACCTTGACGGGCCAGATGTG	420
Db		121	AlaaparglylvaisertlnasnTyrrproillevalGlnasnleuglnlyglmetVal	140
OY		421	CACCAAGGCATCAGCCCCCGCGACCCCTGMAACGGCTGGTGTGAAGTATCGAGAGAAAGCC	480
Db		141	HisglnalatliserProatrgrThreusnnaIatrpvalvsvallleglucglunysala	160
OY		481	TTTAGCCCCGAGGTATCCCATGYTTCACCGGCTTGAGCGAGGGCGCCACCCCGCAGAC	540
Db		161	PheserProgluvalliePrometethetrhalaleuSERgluclyAlathrProglinsap	180
OY		541	CTGAACACGATGTTGAACAACCGTGGGGCGGCAACAGCGCCGATCAGATGCTGAAGAC	600
Db		181	LeuaenthmetleuasnthrvaliglyghisglnaalameGlmetleuylsap	200
OY		601	ACCATCAACGAGAGAGCCGCGAGTGGACCGGCTGTGCACCCCTGTGAGCGCGCCCGTG	660
Db		201	ThrlleaanggluglnalalagluntpraspalguehIsprovalhIsalaglpyrolle	220
OY	.	661	GCCCCCGGCAAGATCGGCGACCCCGCGGCGACGACATGCGCGGGCGCAACGACCCGTG	720
Db		221	AlaprrgIyglImetarargluProrarglyserapllaalaglyThrtsserleu	240
OY		721	CAGAGCAGATCGCCTGTGATGACACGAAACCCCGGTCGCGTGGGCGACATTCACAAG	780
Db		241	GlngluglnllealatrpmethctrhlryasnpropovalProvalGlyasplletrylys	260
OY		781	CGGTSGATCATCTCGGGCTTGAAACAAGATCGTGGGATGTACAGCCCGGTGAGCATCCTG	840
Db		261	ArgrtprrlelleuglyleuasnllysllevalArgmetTyserProvalserlleuen	280
OY		841	GACATCGGCAAGGGCCCCAAGAGGCCCTTCGCGCATCGTGGACCGGCTTCCAAGACC	900
Db		281	AspllleysglnglyPrrplyselutProphetahsparyIrvaaAsparPhetheysthr	300
OY		901	CTGCGCGCGAGCAGGCCACCCAGGACGTGAAGAATGATGATACCGGACGCTGTGTGTG	960
Db		301	LeuaragalaglugglnalalthrdlnmarvalllysaantpmethThrasphrtheueuala	320
OY		961	CAGAACCGCAACCCCGACTGCAAGACCATCTGCGCGCTTCGCGCCCGGCGCCACCTG	1020
Db		321	GlnasnlaaanProaspCyseylstrlleueadgalaleuglyProdyllaaserleu	340
OY		1021	GAGAGATGATGACCGGCGTCCAGGGCGGTGGCGGCGCCGCGCAAGGCCCGCGTGTG	1080
Db		341	GlnglumetmethThalacysglndlyvalglytyrProserhlyvalalargvallieu	360
OY		1081	GCCGAGCGGATGAGCCAGGCGCAACGCGTGAACATCATGATCSADAGSCAATTCACAAG	1140
Db		361	Alaeglualametserclnthrsanser---ThrlleuemecGlnargerxsnheyls	379
OY		1141	GGCCCCGGCGCAAGTCAAGTgcTTCAAATGcGGCAAGAGAGGGCCACATGCGCAAGAC	1200

Db 380 GILPRLYSAGTILEVALYSCYSPHENYCSGLYLVSGILGTYHISTILEALYSAEN 399

QY 1201 TCCCGCGCCCCCGCAGAGGGCTCTGGAAATGTCGCGCAGAGAGGCGCACCAAGATGAAG 1260

Db 400 CYSAAGALAPROAIRGLYLYBSGLCYSTRIPYSCYSGILYLGILGILYHISGIMETLYS 419

QY 1261 GACTGCACCGAGCGCCGAGGCCCAACTTCTGGGCAGAGATGTGCGCCAGCCAGACAGAGGCGCGC 1320

Db 420 ASPTSTHGLAIRGLINLAISNPHENGLIYLSILETRPROSERHISLYSGIYARX 439

QY 1321 CCGCGCACTCTCTGACAGAACCGGACGAGCGCCCGCCGCCACCTGGCCACCGCGCCC 1380

Db 440 PROGLYANPHLEUGLINSERTRPROGU-----PROTHRALAPRO 453

QY 1381 CCGCGGAGAGCTTCCGCTTCGAGAGACACCCCGCCCGCCAGAGAGAGCCCAAGAC 1440

Db 454 PROHLAGUSERPHAIRGHEDILUHTHRMPROHIALAUGLINDILNLYPROLYSAP 473

QY 1441 CGCGAGCCCTTACCGGAGCCCTGACCGCCCTGGCGAGCTGTTCGGACGCGGCCCCCTG 1500

Db 474 -----ArgGLuProLeuThrSerLeuArgSerLeuPheGlySerAspProLeu 489

QY 1501 AGCGAG 1506

Db 490 SERGLN 491

RESULT 13

US-10-339-217-143

/ Sequence 143, Application US/10339217

/ Publication No. US20030198648A1

/ GENERAL INFORMATION:

/ APPLICANT: Buechter, Douglas

/ APPLICANT: Hou, Xiaohong

/ APPLICANT: Marlor, Christopher W.

/ APPLICANT: Rice, William G.

/ APPLICANT: Yang, Wengang

/ TITLE OF INVENTION: Methods for Identifying Compounds which Inhibit Binding of

/ TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA

/ FILE REFERENCE: 111021.143 (ACH-US1)

/ CURRENT APPLICATION NUMBER: US/10/339,217

/ CURRENT FILING DATE: 2003-01-09

/ PRIOR APPLICATION NUMBER: US 60/347,369

/ PRIOR FILING DATE: 2002-01-11

/ NUMBER OF SEQ ID NOS: 144

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 143

/ LENGTH: 491

/ TYPE: PRT

/ ORGANISM: HIV-1

US-10-339-217-143

Alignment Scores:

Pred. No.: 4.7e-116 Length: 491

Score: 2413.50 Matches: 456

Percent Similarity: 95.02% Conservative: 21

Best Local Similarity: 90.84% Mismatches: 14

Query Match: 84.98% Indels: 11

DB: 14 Gaps: 3

US-09-475-704A-4 (1-1509) x US-10-339-217-143 (1-491)

QY 1 ATGGGCGCGCGCGCAGCATCTCTGGCGCGCGAGACTGGACAGTGGGAGAAATCGCG 60

Db 1 MetGtIaIaTgAlaTgAlaSerIleLeuArTgLyGlyLVsLeuAspLyTPrGILuYsIleAg 20

QY 61 CTGGCGCCCGCGCGCAGACACTATGCTGAAGACACTTGSTGTGGCGCCGCGAG 120

Db 21 LeuArGPrGILyGLyLVsLeuYHISerTLeuTleYSHISLeuValTPrAlaSerArGILu 40

QY 121 CTGAGAGGAGCTTGCCTTGAACCCCGGCTCTGCGAGACCGCGGAGGCTGCAAGCATC 180

Db 41 LeuGILuArGPrHeAlaLeuAsnPrGILyLeuLeuGILuTrHrSerGILuGILyCyLeVSGILNle 60



```

QY 181 ATGAAGCAGCTGCGCCGCTGACAGACCGGACCGAGAGCTGGCAGCCTGTACAC 240
   |||
Db 61 IleuysglnleuGlnProAlaIleuGlnThrGlyIleuGlnleuysSerIleuHis 80
QY 241 ACCGTGGCCACCTTACTGCTGTGCAACCGCGCATCGAGTCCGCGACACCAAGAGCC 300
   |||
Db 81 ThrValAlaThrLeuTyrcysValHisAlaGlyIleGluIleArgAspThrIysGluAla 100
QY 301 CTGGACAGATTCAGAGAGAGAGCAACAAGTCCCGACAGAAAGCCAGAGCCCAAGAG 360
   |||
Db 101 leuAspLysIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGln 120
QY 361 GCGGACGCGAGGTGAGCAGAACTACCCCATGTGACAGAACTGAGAGGCGCAGATG 420
   |||
Db 121 AlaAspGlyLysValSerGlnIleAsnTyrrProIleValGlnAsnIleuGlnIleMetVal 140
QY 421 CACCAAGCCATAGACCCCGCACACCTGAACGCTGGGTGAAGGTGATCGAGAGAGCC 480
   |||
Db 141 HisGlnAlaLeuSerProArgThrIleuAsnAlaTrpValIleGlnIleLysAla 160
QY 481 TTCAGCGCCGAGGTATCCCATGTTCAACGCGCTGAGAGAGGCGCCACCCCGACAGAC 540
   |||
Db 161 PheSerProGlnIleIleProMetPheIleAlaLeuSerGlnGlyAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTGAACACCGTGGCGGCGCACAGGCGCCATGCAAGTCTGAAGGAC 600
   |||
Db 181 leuAsnThrMetIleuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetIleuLysAsp 200
QY 601 ACCATCAACGAGAGAGCGCCGACAGTGGAGCGCCTGTGACCCCGGTGAGCGCCCGCTG 660
   |||
Db 201 ThrIleAsnGlnIleGlnAlaIleGlnIleGlnIleGlnIleGlnIleGlnIleGln 220
QY 661 GCGCCGCGCAGATGGCGGACCCCGCGGACGACATGCGCGCGCGCCACAGACCCCTG 720
   |||
Db 221 AlaProGlyIleMetArgIleuProArgIleGlySerAspIleAlaGlyThrIleSerThrIleu 240
QY 721 CAGGACGAGATCGCCTGATGACACAGAACCCCGCTGCGCGGCGCAGCATTTACAAG 780
   |||
Db 241 GlnGlnGlnIleAlaIleTrpMetThrGlyAsnProProValProValGlyIleIleTyrs 260
QY 781 CGGTGATCATCTCTGGGCTGTAAACAAGATCTGCGGATTTACACCCCGCTGACATCTCTG 840
   |||
Db 261 ArgTrpIleIleLeuGlnIleuAsnLysIleValArgMetTySerProValSerIleLeu 280
QY 841 GACATCCGCGCAGGCGCCCAAGAGCCCTCCGCACTAGCTGACCGCTTCTTCAAGACC 900
   |||
Db 281 AspIleLysGlnGlyProLysGlnProPheArgAspTyValAlaAspArgPheIleuThr 300
QY 901 CTGGCGCGCGACGAGCCACACCGAGACCTGAAGAACTGATGACCGAGACCTGCTGCTG 960
   |||
Db 301 leuATGAlaGlnIleGlnIleThrGlnAspValLysAsnTrpMetThrAspThrIleuVal 320
QY 961 CAGAACCCCAACCCCGCATGCAAGACATCTGCGCGCTCTGGCGCCGCGCCACCTG 1020
   |||
Db 321 GlnAsnAlaAsnProAspLysThrIleLeuArgAlaIleuGlnIleProGlyLaseIleu 340
QY 1021 GAGGAGATGATGACCGCTGACAGGCGCGGCGCGCGCCGACCAAGGCGCGCTGCTG 1080
   |||
Db 341 GlnGlnIleuMetThrAlaCysGlnGlyValGlyGlyProSerHisLysValAlaTrpValLeu 360
QY 1081 GCGGAGCGCATGAGCCAGCCCAACCGCTGAACATCATGATGACAGAAAGCAATTTCAAG 1140
   |||
Db 361 AlaGlnAlaMetSerGlnThrAsnAsn---SerIleLeuMetGlnArgSerAsnIleLys 379
QY 1141 GCGCCCGCGCGCAACGTCATGCTTCAACTCGCGGACAGAGAGCGCCACATCGCCCAAGAC 1200
   |||
Db 380 GlyPheLysArgThrValLysCysPheAsnCysGlyLysGlnGlyHisIleAlaArgAsn 399
QY 1201 TGGCGCGCGCGCGCAAGAGGCGCTGGAAGTGGCGCAAGAGGCGCCACAGATGAAG 1260
   |||
Db 400 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGlyHisGlnIleLys 419

```

```

QY 1261 GACTGCACCGAGCGCCGACCACTTCTGGGCAAGATCTGGCCGACGACCAAGAGCGCC 1320
   |||
Db 420 AspCysTrpGlnArgGlnAlaAsnPheIleuGlyIleTrpProSerHisLysGlyArg 439
QY 1321 CCGGCGCACTTCTGACAGAACCGGAGCCGCGCGCCCGCCACCGTGCACCGCCCCC 1380
   |||
Db 440 ProGlnAsnPheIleuGlnAlaAsnArgProGlu-----ProThrAlaPro 453
QY 1381 CCGCGCGAGCTTCCGCTTGGAGAGACCAACCCCGCGCCCAAGAGAGCCCAAGAGAC 1440
   |||
Db 454 ProAlaGlnSerPheArgPheGlnGlnIleThrThrProAlaLeuLysGlnGlnIleLysAsp 473
QY 1441 GCGGAGCCCTTACCGGAGCCCGCTGACCGCGCTGCGCAGCCTGTTCCGCGAGCGCCCTG 1500
   |||
Db 474 -----ArgGlnProLeuThrSerIleuLysSerIleuPheGlnLysAspProLeu 489
QY 1501 AGCCAG 1506
   |||
Db 490 SerGln 491

RESULT 14
US-10-339-217-105
; Sequence 105, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; FILE REFERENCE: 111021.143 (ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 508
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-105

Alignment Scores:
Pred. No.: 1.08e-114 Length: 508
Score: 2387.00 Matches: 454
Percent Similarity: 92.19% Conservative: 18
Best Local Similarity: 88.67% Mismatches: 26
Query Match: 84.05% Indels: 14
DB: 14 Gaps: 3

US-09-475-704a-4 (1-1509) x US-10-339-217-105 (1-508)
QY 1 ATGGCGCGCCGCGCAGCATCTCGCGGCGAGAGAGTGAACAAGTGGAGAAATCCGC 60
   |||
Db 1 MetGlyAlaAlaGlnAlaSerIleLeuArgIleThrIysIleuAspAlaTrpGlnLysIleArg 20
QY 61 CTGGCGCGCGCGCAGAACACATCACTGTGAAGCACTGTGTGGCGCCAGCGCGAG 120
   |||
Db 21 leuArgProGlyGlyLysLysIleTyMetIleuLysHisLeuValITrpAlaSerArgGlu 40
QY 121 CTGAGGCGCTTCCGCTTGAACCCCGCGCTGTGAGACCGCGGAGGCTGCAAGCAATC 180
   |||
Db 41 leuGlnArgPheAlaLeuAsnProGlyLeuIleuGlnIleThrSerGlnGlyCysLysGlnIle 60
QY 181 ATGAAGCAGCTGAGCCCGCTGACAGACCGGACGAGAGAGTGGCGAGCCTGTACAC 240
   |||
Db 61 MetLysGlnIleuHisProAlaIleuGlnIleThrGlyThrGlnIleuLysSerIleuTyAsn 80
QY 241 ACCGTGGCACCTGTACTGTGTCACGCGCGCATGAGGTCCGCAACCAAGAGAGCC 300
   |||

```



```

Db      81 ThrValAlaThrLeuTyrCysValHisGluAsnIleValValArgAspThrLysGluAla 100
QY      301 CTGGACAAGATCGAGAGGAGGAGAAC-----AAGTCCCAAGAGAAAGACCCGACGGCC 354
Db      101 LeuAspLysIleGluGluGluGlnAsnLysIleLysSerGlnGlnLysThrGlnGlnAla 120
QY      355 AAGGAGCCGAGCGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAG 414
Db      121 LysAlaAlaAspGluLysValSerGlnAsnThrProIleValGlnAsnLysGlnGln 140
QY      415 ATGTGTACACCGAGCCATCAGCCCGCCGACCTTGACGCTGGTGAAGTGTACGAGAG 474
Db      141 MetValHisGlnAsnLysSerProArgThrLeuAsnAlaThrValLysValIleGluGln 160
QY      475 AAGGCTTTCAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGGAGGGCCGACCCCC 534
Db      161 LysAlaPheSerProGluValIleProMetPheThrAlaLysSerGlnGluAlaThrPro 180
QY      535 CAGGACCTGAACACGATGTTGAACACGTTGGCGGCGGACACGAGCCGCAATGAGATGCTG 594
Db      181 GlnAspLysSerThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeu 200
QY      595 AAGGACACCATCAACGAGAGAGCGCGGAGTGGAGACCGCTGCACCCCGTGCAGGCGCG 654
Db      201 LysAspThrIleGluGluGluAlaIleGluLysProAspArgLysIleProValHisAlaGly 220
QY      655 CCCGTGGCCCCCGGCGGATGCGGACCCCGCGGCGGACGACATCGCCGCGGCGCACGAC 714
Db      221 ProMetAlaProGlyIleLeuAspGluProArgLysSerAspIleAlaGlyThrThrSer 240
QY      715 ACCCTGCAGAGAGATCGCTGATGACGACGACCCCGCGTGGCGGCGGCGGCGGACATC 774
Db      241 ThrLeuArgGluGlnIleAlaThrMetThrSerAsnProProIleProValGlyAspIle 260
QY      775 TACAAACGGTGAATCATCTCTGGGCTGAACAAGATCGTGGATGTAACCGCTTCTTC 834
Db      261 TyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetCysSerProValSer 280
QY      835 ATCTTGGAACATCGGCGGAGGCGCCCAAGAGACCTTTCGCGACATACTGACCGCTTCTTC 894
Db      281 IleLeuAspIleArgGlnGlyProLysGluProPheArgAspLysValAspArgPhePhe 300
QY      895 AAGACCTGTGCGCGCGGAGGCGGACCGGACGTAAGAACTGATGACCGGACCGGACCG 954
Db      301 LysAlaLeuArgAlaGluGlnAlaThrGlnAspValLysAsnThrPheThrGluThrLeu 320
QY      955 CTGTGTGCAGAACGCAACCCCGACTGCAGAACCATCTCGCGCTTCGCGCCCGCGCGCC 1014
Db      321 LeuValGlnAsnAlaAsnProAspArgLysThrIleLeuLysAlaLeuGlyIleGlyAla 340
QY      1015 ACCCTGAGAGATGATGACCGCTTCGCGGCGTGGGCGGCGCCGCGCAAGGCGCCG 1074
Db      341 ThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyProSerHisLysAlaArg 360
QY      1075 GTGCTGCGCGAGCGGCGTGAAGCGGCAACGCTGATCATCTGATGAGAAAGACCAAC 1134
Db      361 ValLeuAlaGlnAlaMetSerGlnAlaAsnMetThrAsnIleMetMetGlnLysSerAsn 380
QY      1135 TTCAAAGGCGCGCGCGGCAACGTCATGCTTCAACTGCGGCAAGAGGCGCATCGCC 1194
Db      381 PheLysSerSerLysArgIleValLysCysSerAsnCysGlyLysGluGlyHisIleAla 400
QY      1195 AAGAACTGCGCGCGCGCGCGCAAGAGGCTGTGAAAGTGTGCGGCAAGAGGCGCAAC 1254
Db      401 ArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGln 420
QY      1255 ATGAAAGATGCAACGAGCGGCGGCGCAACTTCCTGGGCAAGATCTGGCGCGGCGCAAG 1314
Db      421 MetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLys 440
QY      1315 GGGCGCGCGCGCAACTTCTGCAGAAACGCGAGCGGCGCGCGCGCGCGCGCGCGCG 1368
Db      441 GlyArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSer 460

```

```

QY      1369 -----CCCAACCGCGCGCGCGCGGAGAGCTTCGCTTCGAGAGAAC 1410
Db      461 PheArgAsnArgProGluThrProThrAlaProProAlaGluSerPheArgPheGluThr 480
QY      1411 ACCCGCGCGCGCGGAGAGCGGCAAGAGCCCGGAGCCCTTACCGGAGCGCGCTGACCGCC 1470
Db      481 ThrProThrProLysGlnGluProLysAsp-----ArgAspProLeuThrSer 496
QY      1471 CTGCGGAGCTGTTCGCGAGCGGCGCGCGCTGAGCCAG 1506
Db      497 LeuLysSerLeuPheGlySerAspProSerSerGln 508

RESULT 15
US-10-339-217-104
; Sequence 104, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for identifying compounds which inhibit binding of
; FILE REFERENCE: 111021.143 (ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FASTA3.0 for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 491
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-104

Alignment Scores:
Pred. No.: 1,29e-114 Length: 491
Score: 2385.50 Matches: 453
Percent Similarity: 93.82% Conservative: 18
Best Local Similarity: 90.24% Mismatches: 20
Query Match: 84.00% Indels: 11
DB: 14 Gaps: 3

US-09-475-704A-4 (1-1509) x US-10-339-217-104 (1-491)
QY      1 ATGGGCGCGCGCGCGGAGGATCTCTGGCGGCGGAGAGCTGACAACTGGAGAAATCCGC 60
Db      1 MetGlyAlaArgAlaSerIleLeuArgGlyGluLysLeuAspLysTrpGluLysIleArg 20
QY      61 CTGGCGCGCGCGGAGGAGCACTAATGCTGAAGACCTGTGTGGGCGGCGCGCGAG 120
Db      21 LeuArgProGlyLysArgLysHisIleMetLeuLysHisLeuValTrpAlaSerGlnIle 40
QY      121 CTGAAGGCTTGGCCCTGAACCCCGGCGCTGTGAAGACCGCGGAGGCTGCAAGAGATC 180
Db      41 LeuGluArgPheAlaLeuAsnProGlyLeuGluGluThrSerGlnGlyCysLysGlnIle 60
QY      181 ATGAAGACGCTCAAGCCCGCGCTTCGACAGCCGCGACCGGAGAGCTGCGAGCTGTACAC 240
Db      61 IleLysGlnLeuHisProAlaLeuLysThrGlyThrGluGluLeuArgSerLeuTyrAsn 80
QY      241 ACCGCGGCGGCGGCTGTACGCTGTGACGCGGCGGAGTCCGCGGACCAAGAGGCGCC 300
Db      81 ThrValAlaThrLeuTyrCysValHisGluAsnIleGluValArgAspThrLysGluAla 100
QY      301 CTGACAAGATCGAGAGGAGGAGAACTACCCCATCGTGCAGAACTGCAGGGCCAG 360
Db      101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaLysAla 120
QY      361 GCGGACGCGGAGGTGAGCGGAGAACTACCCCATCTGTGCAAGAACTGAGGCGGCAATGTGTG 420

```

```
Db      121 AlaArgGluGlyValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
Qy      421 CACCAAGGCATCAGCCCGCCGACCCCTGAAGCCTGGGTGAAGTGAATCGAGAGAGGCC 480
Db      141 HisGlnAlaIleSerProArgThrLeuAsnAlaIleValValIleGlnGluLysAla 160
Qy      481 TTCAGCCCGGAGTGTATCCCATGTTCAACGCCCTTACGAGGAGGAGGCCACCCCGCAGAC 540
Db      161 PheSerProGluValIleProMetPheThrAlaLeuSerGlnGlyAlaIleThrProGlnAsp 180
Qy      541 CTGAACACGATGTTGAACAACGCTGGGCGGCAACAGCCGCCATGCAATGCTGAAGAC 600
Db      181 LeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaIleMetGlnMetLeuLysAsp 200
Qy      601 ACCATCAACAGAGAGGCCCGCCAGTGGACCCGCTGCACCCCGTGCAGGCCCGCCGCTG 660
Db      201 ThrIleAsnGluGlnAlaIleArgIleTyrAspArgLeuHisProValHisAlaGlyProAla 220
Qy      661 GCGCCCGGCGAGATGGCGGACCCCGCGGAGGACATGCGCGGCGGCAACAGACCCCTG 720
Db      221 AlaProGlyGlnMetArgGlnTyrProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
Qy      721 CAGAGACAGATCGCTGATGAACAACAACCCCGCTGCGCGTGGCGGACATCTAACAAG 780
Db      241 GlnGluGlnIleAlaIleThrMetThrGlyAsnProProValProValGlyAspIleTyrLys 260
Qy      781 CGGTGATATATCTGGGCGCTGAACAAGATCGTGGCGGATGATACAGCCCGCTGAGCATCTG 840
Db      261 ArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
Qy      841 GACATCCGCGCAGAGGCCCAAGAGACCTTCCGACTACGTGAGACCGCTTCTTCAAGACC 900
Db      281 AspIleIleLysGlnGlyProLysGlnTyrProPheArgAspTyrValAspArgPhePheLysVal 300
Qy      901 CTGCGCGCGGAGAGGCCCAAGAGACCTTCCGACTACGTGAGACCGCTTCTTCAAGACC 960
Db      301 LeuArgAlaGluGlnAlaIleThrGlnAspValLysAsnTyrMetThrAspThrLeuLeuIle 320
Qy      961 CAGAAAGCCCAACCCCGACTGCAAGACCACTCGCGGCTCTCGGCGCGCGCCACCTG 1020
Db      321 GlnAsnAlaAsnProAspLysThrIleLeuLysAlaLeuGlyProAlaAlaSerLeu 340
Qy      1021 GAGGAGATGATACCGCTGCAAGGCGGTGGCGGCGCGCCGCAACAAGGCCCGCTGCTG 1080
Db      341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgValLeu 360
Qy      1081 GCCGAGGCGATGAGCCGAGCCCAACAGCGTGAACATCATGATGACAGAAAGCACTTCAAG 1140
Db      361 AlaGlnAlaMetSerGlnAlaAsnSer---AsnIleMetMetGlnArgSerAsnPheLys 379
Qy      1141 GGGCCCCGGGCGCAAGCTCAAGTCTTCAACTCGCGGCAAGAGAGGCCCACTCCGCAAGAAC 1200
Db      380 GlySerLysArgIleValLysCysPheAsnCysGlyLysGlnGlyHisIleAlaArgAsn 399
Qy      1201 TSCCGGCGCGCCCGCAAGAGGCGTGTGGAAGTGGCGGCAAGAGAGGCCCAAGAGCGCG 1260
Db      400 CysArgAlaProArgLysLysGlyCysTyrLysCysGlyGlnGlnGlyHisGlnMetLys 419
Qy      1261 GACTGCACCGAGCGCGCCCAACTTCTGGGCAAGATGTGCGCCCAAGAGAGCGCGC 1320
Db      420 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTyrProSerHisLysGlyArg 439
Qy      1321 CCGGCAACTTCTGCAAGAACCGGACAGGCGCGCGCCCGCCCAACGCTGCGCCACCGCCCC 1380
Db      440 ProGlyAsnPheLeuGlnIleSerArgProGlu-----ProThrAlaPro 453
Qy      1381 CCGGCGAGAGCTTCGCTTCAAGAGAACACCCCGCGCCCAAGAGAGGCCCAAGAGAC 1440
Db      454 ProAlaGluSerPheArgPheGlnGluGlnThrThrProAlaProLysGlnGlnSerLysAsp 473
Qy      1441 CGCGAGCCCTTACCGCGAGCCCTTGAACGCGCTGCGGAGCCTGTGCGAGCGGCGCCCTG 1500
```

Db 474 -----ArgGluProLeuIleSerLeuLysSerLeuPheGlySerAspProSer 489

Qy 1501 AGCCAG 1506

Db 490 SerGln 491

Search completed: September 1, 2005, 23:51:13  
Job time : 408.088 secs